1978-06

A non-parametric analysis of the HRM data

Carter, William L.; Yeatman, Lawrence L.

A NON-PARAMETRIC ANALYSIS OF THE HRM DATA

William L. Carter

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A NON-PARAMETRIC ANALYSIS OF THE HRM DATA

by

William L. Carter

Lawrence L. Yeatman

June 1978

Thesis Advisor: Cdr. C. B. Gustafson

Approved for public release; distribution unlimited

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) A Non-parametric Analysis of the HRM Data	5. TYPE OF REPORT & PERIOD COVERED Master's Thesis June 1978
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) William L. Carter Lawrence L. Yeatman	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	12. REPORT DATE June 1978
	13. NUMBER OF PAGES 127
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Survey of Organizations
Survey Guided Development
HRM Survey
Trend Analysis
Non-Parametric Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This thesis provides a statistical analysis of historical HRM data in an attempt to determine if significant positive trends exist. Non-parametric statistical techniques using rank-ordering concepts were used for the analysis.

Analysis failed to show the trends predicted by the literature on survey-guided development. No significant overall trends were detected by any of the statistical tests used. A conclusion that no significant changes have occurred in Navy

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

attitudes as the result of efforts by the HRMSS was considered suspect. The authors discussed the possible explanations for the discrepency between the conceptual prediction and the observed results.

In conclusion, the analysis was not able to reject the null hypothesis. Possible areas of future research are presented for consideration by interested parties.

A NON-PARAMETRIC ANALYSIS OF THE HRM DATA

by

William L. Carter
Lieutenant, United States Navy
A. A., University of Florida, 1969, B. S., Livingston University, 1971

Lawrence L. Yeatman
Lieutenant, United States Navy
B. A., University of Texas, 1971

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
June 1978

ABSTRACT

This thesis provides a statistical analysis of historical HRM data in an attempt to determine if significant positive trends exist. Non-parametric statistical techniques using rank-ordering concepts were used for the analysis.

Analysis failed to show the trends predicted by the literature on survey-guided development. No significant overall trends were detected by any of the statistical tests used. A conclusion that no significant changes have occurred in Navy attitudes as the result of efforts by the HRMSS was considered suspect. The authors discussed the possible explanations for the discrepency between the conceptual prediction and the observed results.

In conclusion, the analysis was not able to reject the null hypothesis. Possible areas of future research are presented for consideration by interested parties.

TABLE OF CONTENTS

LIST OF FIGURES

I. HISTORY OF THE HUMAN RESOURCE MANAGEMENT (HRM) PROGRAM

The initial thrust for, what is now termed, the Human Resource Management Support System (HRMSS) was originally provided by Admiral Elmo R. Zumwalt, Jr. In October 1970, the then Chief of Naval Operations instituted a new policy-making and planning section(Pers-P) in the Bureau of Naval Personnel. The dual responsibility of Pers-P (to become known as the "People Program" Code) was to promote the "worth and dignity" of all naval personnel and increase organizational effectiveness. Its immediate goal was to communicate directly with people at all levels of the organization, from the commanding officer to the recruit, on problems relating to the human condition. Its ultimate goal was to establish new people oriented programs and policies.

The Human Resource Management (HRM) Project was established in January 1971 and became part of the Human Resource Management Project Office (Pers-Pc), when it was created in March 1971. The mission of the HRM project was to improve the management of the Navy's human resources by enhancing the understanding of and communications with its people. Personnel assigned to the project, the so-called Z-55 group, were chartered to develop and evaluate new ideas in the human relations area. The group derived two basic objectives from its mission. (1) test and evaluate practical applications of knowledged science in the Navy with respect to assisting naval units to effectivelly accomplish their mission and (2) to determine the most appropriate methods or techniques of conveying the successes of these findings and applications throughout the Naval

Establishment. Initial efforts were aimed toward formulating an organization development program because of its practical significance in accomplishing the group's mission. These efforts involved (1) acquiring a better understanding of the problems faced by top management (not just those of individual Commanding Officers (COs)) in trying to resolve the Fleet's human relations problems, and (2) at the CO level, gaining more insight into general leadership styles.

In December 1971, the Command Development Program was established within Pers-Pc as the result of the exploratory work of HRM project personnel. The expressed goal of the new program was to develop leadership and managerial skills at the local level to effectively integrate men and mission. As a step in this direction, "N-Man" typologies (i.e., descriptions of primary leadership styles) were generated and used in "Step One" seminars. In these introductory seminars, designated commands were introduced to the N-Man concept as a technique for diagnosing their own leadership styles and, most importantly, for determining the possible impact various styles had on the management of human resources. A seven-step organizational development package, based on behavioral science concepts and principles, was offered to the commands as part of a pilot program.

To attack more specific social and human relations problems, a number of problem-solving options, under the general sponsorship of Pers-Pc, were also offered to interested commands. These options included (1) managerial seminars and workshops, (2) treatment centers for chronic alcohol and drug abusers (Long Beach and Miramar, Ca, respectively) and (3) an intercultural relations center for educational assistance in that area (Coronado, California). A similar Department of Defense sponsored educational center in the area of racial relations (Patrick Air Force Base,

Florida) was also started.

In April 1972, Pers-Pc was redesignated as the Human Resource Development (HRD) Project. The HRD Project encompassed: (1) command development, (2) race relations, (3) intercultural relations, and (4) drug and alcohol abuse education, rehabilitation, and control. Human Resource Development Centers (HRDCs) were subsequently established in Newport, San Diego, Pearl Harbor, and Norfolk to aid in the diagnosis and solution of problems in each area. Actual HRDC field work was conducted by Navy trained in-house consultants who responded independently to "requests for assistance" in their area of cognizance.

The survey-guided development phase began in August 1973, when Pers-Pc distributed the first draft of the Navy Human Goals Plan. This plan envisioned that specialists would (1) use a comprehensive survey approach to guide their activities, operating through Human Resource Development Centers (later to become known as Management Centers) and Detachments and (2) begin to function as a single multipurpose team, with the total ship or unit as their target instead of circumscribed problems. Implementation of this plan subsequently began, after pilot testing, on a fairly limited basis. This marked the initial beginnings of the Human Resource Management (HRM) Program in its current form.

In January 1974, the HRDCs and HRDDs were transferred from Pers-P to the Commander-in-Chiefs of the various Fleet commands for full implementation and whereupon their names were changed to Human Resource Management Centers and Detachments (HRMCs and HRMDs). However, actual funding shifts were not accomplished until 1 July 1974. With these shifts, the HRM Program was to be applied to all commands, both fleet and shore establishments. The practical

apparatus for implementing the HRM Program was termed the Human Resource Management Cycle. This cycle, in its broad initial outlines, involved (1) administration of the survey itself, a session where consultants provided feedback on command deficiencies and a 5-day dedicated period during which HRM consultants implimented practical remedies, and (2) a 6 to 10 month follow-up designed to assist the command as required.

Conceptually, the HRM Program differed from its Pers-P predecessor in two ways.

1. Equal opportunity replaced race relations (i.e., Black-White relations) as a guiding concept in order to accommodate the needs of all minorities, including women.

2. The notion of overseas diplomacy, predicated on the role of naval personnel as goodwill ambassadors, augmented the simpler notion of intercultural relations, which had focused primarily on the adjustment process of overseas personnel and their dependents.

Figure 1 traces the evolution of the HRM Program. (Wilcove, 1976)

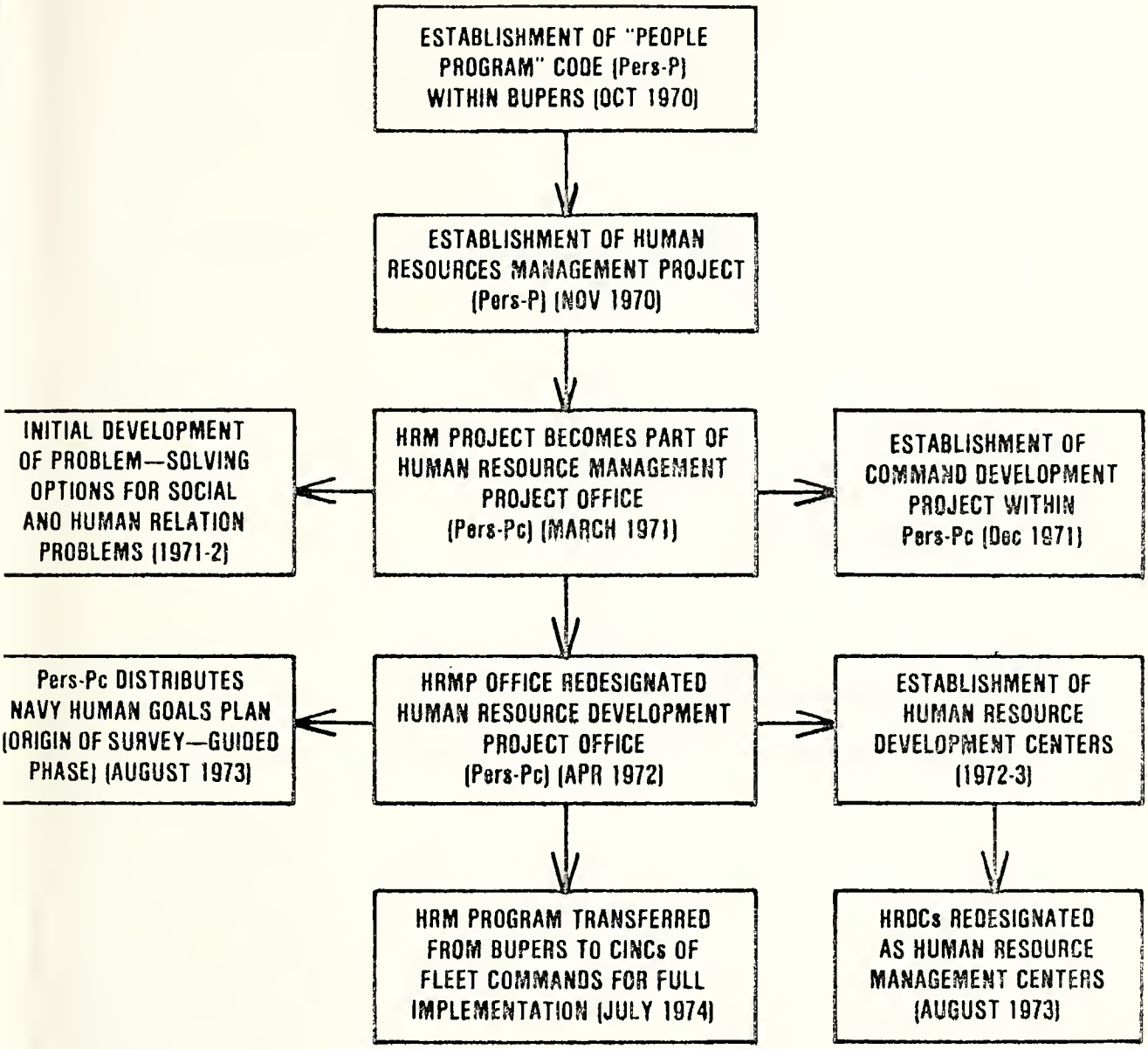

Figure 1 - FLOW CHARTING THE EVOLUTION OF THE NAVY'S HUMAN RESOURCE MANAGEMENT PROGRAM

II. AN OVERVIEW OF SURVEY-GUIDED DEVELOPMENT

A. DEFINITION OF SURVEY-GUIDED DEVELOPMENT

This chapter is designed to answer the question, "What is Survey-Guided Development?" It is an important question since some people may want a brief and simple answer, others a detailed explanation, and still others something in between.

What is Survey-Guided Development? The name itself suggests two important facets of the answer: First, it is the development or improvement of the capacity of organization members to function effectively in accomplishing their day-to-day and longer-range tasks. In a truly effective organization, people at all levels of the organizational hierarchy are motivated and able to solve problems in a way which anticipates and eliminates problems without creating new ones. Second, Survey-Guided Development utilizes instrumentation which asks for many people's perceptions about important aspects of organizational functioning. The most significant of these aspects are those that affect performance, satisfaction, and motivation in the organization. By comparing people's perceptions of how the organization does function with some standards of how it might or should function, the areas needing development become identifiable.

B. DIAGONSTIC INFORMATION

An organization is not a fixed entity located in an unchanging environment. Rather, it is in constant flux with its members attempting to adapt to changing conditions both inside and outside the organization. Constructive adaptation to these changing conditions necessitates having information about how the organization is working and where there are problems in its functioning. Survey-Guided Development utilizes instrumentation to efficiently provide adequate and accurate information about how the organization functions. This information provides organization members with the kind of descriptive picture which will help them to monitor and adapt to changing demands and situations more smoothly and to anticipate and solve problems more effectively.

To discuss these two points in any more detail it is important to understand the theory and assumptions underlying Survey-Guided Development.

Survey-Guided Development is based upon the theory that three factors need to be taken into account in an organizational development effort: the behavior(s) which are problematic, the conditions which create those behaviors, and the conditions creating the problems. These three factors may be stated in the form of three relatively simple questions: What is it that people are doing, or not doing, that is a problem, why are they doing or not doing these particular things, and which of a large number of possible interventions or activities would be most likely to solve the problems by focussing upon why the problems exist? It is crucial that these three questions be answered, and answered correctly, since it would not help the organization if the organizational development effort focussed upon changing behaviors which were not the problem behaviors. Nor would it help if a great deal of effort were put in

toward solving the right problems by changing the wrong conditions or using the wrong interventions to change them. Thus, it is important that changes be based upon an accurate diagnosis of problems and their causes.

Yet, the task of formulating such a diagnosis is difficult since change in orgaizations involves taking into account, at one time, the appropriate behaviors and inter-relationships across many people. These people are: (1) in many jobs and roles, (2) at various levels in the organization, (3) at many points in time, and (4) working under varying conditions. How does one collect this much information? In order to get accurate information for such a diagnosis you need to ask many people within the same organization, but in different work situations, to answer the same set of questions. By collecting the replies you gather multiple perceptions of the organization. This practice lessens the chance that the diagnosis will reflect organizational functioning as viewed from a few unique work situations or as distorted by individual biases. In contrast to this large sample, if just a few people were asked to describe the organization, a very limited number of work situations and personal biases would reflect the actual state of the organization's functioning. Think of your own organization: do you know any small group of people who could give you a picture of the organization that would accurately represent its functioning at all levels in all departments regarding all their operations? The larger and the more complex the organization, the more an accurate diagnosis appears to depend upon gathering multiple perceptions.

Once multiple perceptions have been gathered, there is a need to combine and summarize them in a meaningful way so that the amount of information to be digested is not overwhelming.

This summary will tell how an organization is functioning. Yet, on the other hand people are likely to be even more interested in how well their particular organization functions. This means that something is needed with which tc compare the summarized answers; in other words, one needs a standard of comparison. There are at least four kinds of comparisons that could be made: (1) how the organization functions now can be compared with how it functioned at some time in the past; (2) how the organization functions now can be compared with how other organizations of the same type function; (3) how the organization functions now can be compared to how people within the organization would like it to function; or (4) how the organization functions now can be compared to how research and experience have demonstrated it should function. Using the comparisons described in points (1) and (2) above require that one should have answers to the same questions summarized in a similar way for either (1) the same organization at more than one point in time or (2) for more than one organization. In order to make comparisons like those described in points (3) and (4), an estimate of what people want and how they think their organization should function, as well as the accumulated knowledge about effective functioning is needed.

Survey-Guided Development relies upon multiple means for gathering the needed information for assessing how well an organization functions. Interviews, unobtrusive measures, personal impressions and standardized questionnaires to mention a few. It appears that questionnaires are becoming more widely utilized than other means. That is, a machine-scored survey with a standard set of questions shown to be relevant and meaningful in many types of organizations is administered to organization members. This process of using a standardized instrument enables one to make comparisons (1) and (2) listed above. The members'

responses are stored and can be re-called when an organization wants to compare itself with all, or various kinds of, other organizations which have used the survey. (Babbie, 1973) Comparisons such as those described in points (3) and (4) above may be made by including relevant questions on the surveys or by some more informal method such as a meeting where people list their preferences. These preferences and survey questions would then be compared to how the organization is functioning now as described by responses to the standardized questions.

How would one know whether or not he's asking about the most important aspects of organizational functioning? This is a central question since if the survey does not measure the important aspects, then it cannot identify the important problems. The answer, however, is quite straightforward. Aspects of organizational functioning are important when they: (1) affect the performance, motivation, and behavioral reactions of organization members, (2) fit into a coherent scheme or model of how successful organizations function; and (3) are generally known to be crucial to the day-to-day or future survival of the organization.

The questions utilized on the questionnaire are those which have been found to meet all three of the points listed above. Furthermore, only reliable questions (i.e., those measuring stable characteristics not subject to meaningless variation) are included in the standard survey.

C. GOALS OF SURVEY-GUIDED DEVELOPMENT

Thus far, the discussion has focused upon a general definition of Survey-Guided Development and some description of the theory and assumptions underlying this approach. Our

attention will now turn to the more specific kinds of things that Survey-Guided Development is designed to accomplish.

The ultimate goal of Survey-Guided Development is to facilitate interventions or changes in organizational functioning which will lead to increased organizational effectiveness. This is to be accomplished by providing accurate and useful information about how an organization actually functions, how it might ideally function, and how to make the actual functioning more like the ideal functioning.

As written, the ultimate goal of Survey-Guided Development can be expressed in a single statement. There are a number of initial requirements that need to be met, however, if this goal is to be successfully accomplished. First, the organization leaders and members involved in the development effort must have accurate and useful information about how effective organizations function or operate. Second, the organization leaders must decide how their organization should "ideally" function. Third, organization leaders and members must have accurate and useful information about how their organization functions presently and why it functions the way it does. Fourth, organization leaders and members must be aware of existing discrepancies between current organizational functioning and how they would like their organization to function. Fifth, where such discrepancies exist, organization leaders and members must lessen or eliminate them by planning and carrying out a sequence of activities designed to make the organization function more like their "ideal" model. Sixth, after some period of time, the organization leaders and members must have information about the effects of the seguence of activities designed to lessen the discrepancies.

Let's lock at these requirements in more detail. Most

people have some notions or ideas about what makes an organization successful. They may not have intentionally developed these notions nor mentioned them to anyone. Nevertheless, if asked the question, "What makes an organization successful?" most of these people would respond fairly quickly. When these notions are put together in a coherent, meaningful way, we say that a person has a model of effective organizational functioning. When such a model has been tested and found to hold true, it can be said to be a "valid" (i.e., accurate and useful) model. Let us assume that one person thinks that the most effective organizations are run by autocratic leaders while another person thinks that participative leaders are more effective. If two organizations similar in all ways except for the type of leaders employed were studied, and one organization was found to be more effective than the other, the two models of effective leadership would have undergone an initial test. The model of leadership employed by the more effective organization would then be considered the more valid model of the two. Although models of effective organizational functioning include many factors in addition to leadership, this example illustrates one way in which the validity of models is tested.

Survey-Guided Development is based upon a model of effective functioning which has been tested in many kinds of organizations. The more like this model an organization functions, the more effective it tends to be. Organization leaders and members embarking on Survey-Guided Development efforts must understand the model if the effort is to ultimately succeed.

Once the ways in which the most effective organizations function are known, the organization should decide how closely it wants to or can approximate this model; that is, a goal should be set. In development efforts it is

important that organization leaders and members have a model of effective functioning toward which they can strive, since without such a model they would tend to resemble nomads. Think of it this way: When you don't know where you want to go, how do you find the way and how do you know when you've arrived? (Mager, 1968)

Once a goal has been selected, information about how the organization functions now can be collected and compared to how it would function if the goal were achieved. If there is true commitment to the goal, organization leaders and members are motivated to lessen the differences by planning changes and activities which will move the organization closer to that goal. After the changes have had a chance to work, organizational functioning is assessed again to see how much movement actually took place. If there are still discrepancies, or if a new or higher goal is set, additional changes are planned, implemented, and evaluated. It is in this planned, monitored, and controlled manner that the most constructive and lasting improvement occurs.

At this point it is important to gain a perspective of the sequence of steps involved in the development effort. The steps designed to accomplish the goal and meet the six requirements discussed above are summarized in Figure 2.

REQUIREMENTS	STEPS
1. Knowledge of valid model of effective organizational functioning	1. Conceptual Training (i.e., providing the model of effective organizational functioning)
2. Selection of an ideal state of organizational functioning	2. Goal Setting
3. Collection of information about how the organization functions presently	3. Administration of the standardized survey to organization members
4. Recognition of where discrepancies between actual and ideal functioning exist	4. Diagnosis of present organizational functioning
5. Lessening of discrepancies between actual and ideal functioning	5a. Feedback of survey data to the work groups that generated the data (i.e., feedback of work group level data); Planning and implementing action steps at the work group level
	5b. Feedback of survey data about the whole organization to system leaders; Planning and implementing action steps at the system level.
6. Evaluation of the effects of the change activities	6. Re-administration of a standardized survey

Figure 2 - REQUIREMENTS OF SURVEY-GUIDED DEVELOPMENT AND THE STEPS DESIGNED TO MEET THEM

D. SURVEY-GUIDED DEVELOPMENT STRATEGY

Because of the sequential nature of Survey-Guided Development, there are some general guidelines and perspectives which may affect the overall success of the development effort.

Basic to the development process is the theory that the motivation to change is created by the realization that the present state of organizational functioning differs from the ideal model, (i.e., a discrepancy exists between what is desired and what actually exists). However, it does not necessarily follow that the largest discrepancies create the greatest motivation to change. Rather, a moderate discrepancy may be more motivating. A moderate discrepancy often indicates a problem area which needs a significant amount of work, but which is not such a sizeable problem that it could not be solved within a reasonable time period and with available or obtainable resources. In contrast, very small discrepancies are easy to forget about, and huge discrepancies are discouraging. It is the consultant's job to point out and focus on motivating discrepancies.

In a development effort it is crucial that people near the top of the organization are motivated to change and are supportive of changes taking place lower in the organizational hierarchy. The importance of this motivation and support lies in the necessity for creating a climate of constructive change at the top levels which will foster development at all levels below. Deciding how high in the organization one must go depends upon the level at which people have some real power and control over what happens below them in the organization.

One way of building a climate supportive of change at the higher levels is to begin the development activities at those levels, and then soon after they have begun, start the activities at lower levels. This strategy has been found to be more effective than starting at lowest levels of the organization, starting at all levels simultaneously, or dwelling exclusively upon any one level.

In addition, it is useful to know why the organization unit became involved in the development effort in the first place. Was the development effort "laid on"; that is, were people told to use Survey-Guided Development whether or not they wanted to? Is it just one more attempt to be innovative? Was it asked for because of a felt need to examine and solve problems? The answers to these questions and others like them indicate how much time and effort need to be put into establishing what Survey-Guided Development has to offer the particular organization and how likely such and effort is to succeed. If the development effort was forced on people, they may resist it, at least at first, regardless of how much sense the approach makes to them. A fair amount of time and effort might be needed to lessen this resistance. If the organization is trying Survey-Guided Development merely to do something new and different, people may not be truly committed to changing the way they work, even when such changes would lead to constructive improvement. On the other hand, an organization that seeks Survey-Guided Development because it wants assistance in identifying and solving problems is likely to respond quickly and constructively to the information and assistance it receives.

Finally, a few words about what Survey-Guided Development is not. (1) It is not a "shoot-from-the-hip" operation to be handled in a slipshod manner. Rather, it is a systematic program of activities which need to be planned

in advance. (2) It is not a "piece-meal" operation from which bits and pieces get chosen or left out. Rather, it is a coherent, integrated set of activities designed to move the organization closer and closer to its desired level of functioning.

E. THE HRM CYCLE

The Navy Human Goals Plan (Chief of Naval Operations Instruction (OPNAVINST) 5300.6B, 1975) requires at least five common elements for each HRM Cycle, i.e., (1) an initial contact or briefing between the consultant and the commanding officer (CO); (2) administration of the HRM Survey; (3) a meeting between the consultant and CO to feedback the survey data; (4) a dedicated HRM period, Human Resource Availability (HRAV), at the end of which a Command Action Plan (CAP) is written or revised; and (5) a six to ten month follow-up visit. In most cases, points one through four must be completed in a six to eight week period. In other cases, as little as two weeks may be available. It is important to recognize that a Survey-Guided Development effort can fit within these time constraints in more than one way.

This chapter points out that, ideally, Survey-Guided Development activities focus on development at two levels: the system level and the group level. In some situations, however, time and resource constraints will preclude a full Survey-Guided Development effort conducted according to all the guidelines suggested in this chapter. Despite this, there are many possible ways of adapting activities to the basic elements of Survey-Guided Development.

In approaching such an adaptation, certain minimum

components of Survey-Guided Development must be kept in mind. They are: (1) use of the HRM Survey, (2) data analysis by the consultant, (3) feedback of survey data, (4) problem-identification by command members, (5) solution-generation and implementation, and (6) monitoring of effects of solutions. Obviously, these are rather general terms and many of them could take various forms. In strategizing any such adaptation of Survey-Guided Development, it is recommended that in each situation the effort be planned with an ideal set of activities in mind. Where real world conditions impinge on this ideal set, adaptation and compromise will be necessary. Taking this planned approach will result in an effort that more closely approximates the ideal than one that was developed by first identifying all possible constraints in a situation and their potential impact and then designing activities which will not conflict with them. In other words, be constrained only when forced to be.

Careful readers will have noticed that thus far no mention has been made of providing a model of organizational functioning or selecting a goal state of functioning. Survey-Guided Development, in its ideal form, makes these two activities very explicit. Doing so has many benefits for the total effort such as motivating activities and making survey feedback more meaningful. When time and resources do not allow such an explicit model-setting activity as specific training, the consultant must take advantage of any opening he has to provide conceptual inputs. These can occur during several points in an effort. Even on the initial visit with the CO, the consultant may find it useful to share some conceptual information helpful in establishing a frame of reference for the activities to be conducted. During feedback meetings with individuals or groups, the consultant may be able to do some more model-setting when the situation calls for an informational

input of some kind. Beyond these opportunities, it may be necessary to rely on implicit models and goals which command members may already hold and capitalize on those areas where their models seem to complement the model underlying the effort.

III. USING A STANDARDIZED QUESTIONNAIRE

This chapter covers (1) features of a standardized survey and why it is chosen as a core tool for gathering data in Survey-Guided development, and (2) the structure and content of the Human Resource Management Survey which will be referred to throughout subsequent chapters.

A. WHAT IS A STANDARDIZED SURVEY

Survey-Guided Development is based on perceptions of organizational functioning gathered from organization members who respond to a standardized questionnaire. To understand why the survey is an essential tool used throughout the process, we must first consider the characteristics of a standardized survey. Primarily, the term standardized refers to the fact that the survey is comprised of predetermined questions and answer alternatives, uniformly arranged, explained, and administered to all respondents. The questionnaire is devised so that its content is relevant to a large population of people who will use it, and also so that it may be used over a considerable period of time without major revisions. When these conditions are met, the same survey can be administered to many groups of people over time. This possibility of gathering data from large numbers of people allows grouping of all responses together to form norms for the survey, which add to its standardized nature. Briefly, norms are averaged survey responses over large numbers of respondents which can be used as a standard with

which subsequent respondent groups can compare their data. Norms are more fully discussed in Section III of this chapter, but they are mentioned here as an important part of the usefulness of a standardized survey.

The third important characteristic of a standardized survey, particularly one for use in Survey-Guided Development, is that it is closely tied to a theory of organizational functioning. The theory is crucial since it provides a basis for (1) deciding what questions should be included, (2) using the data to accurately describe the state of the organization at the time of the survey administration, and (3) interpreting responses and using the data to improve organizational functioning.

B. WHY USE A STANDARDIZED SURVEY

There are several reasons why a survey is the basis for collecting data for use in the development effort. First, a standardized survey can be efficiently scored and administered to many people. This is important because it allows us to gather multiple perceptions about organization and work group functioning which are useful for assessment and feedback activities in Survey-Guided Development. The more individual perceptions gathered, the more accurately the data describes the aspects of the work situation as it truly operates.

Second, since the survey is tied to a theory of organizational functioning, it measures behaviors and conditions which influence various outcome characteristics (e.g., retention rate, productivity, combat-readiness). This allows looks at aspects of the work situation which partially cause and influence the quality of these outputs,

rather than measuring only the outputs themselves.

Third, survey data is ideally provided by all organization members. Thus, they provide tangible, manageable information which system members themselves can use as a tool for identifying strengths and weaknesses of their command or work group, which caused them to respond to survey questions as they did. The data encourage personnel to look at specific examples of concrete conditions and behaviors which can be improved. They can then plan, on the basis of what their data shows, to work toward and increased level of effective functioning.

C. EFFICIENT SURVEY CHARACTERISTICS

Survey Items. The single survey question is the most basic unit of data and provides the most specific information. In this section, the focus is on considerations which sould be made in deciding how items are to be included on the survey. In taking the intended respondents into account, each question should be clear and relevant to the respondent. Each question should be geared to some concrete aspect of the work situation, and should not be asking about more than one behavior or condition at a time. Answer alternatives should also be clearly relevant to the question, and sufficient in number and content for the respondent to be able to express his opinion adequately. The survey should include an optimal number (neither too few nor too many) of questions to adequately measure the facets of any given area of the work situation. The ability of the question to measure what it is supposed to be measuring, as well as its relevance over time and slight variance in work situations should be taken into account. In this regard, it is sometimes useful to include questions which have, from

previous use, been shown to be effective in measuring particular areas of the work situation. Finally, any item chosen for inclusion on the survey should be justified in terms of its usefulness (1) for research, or to test whether a question which might be valuable is in fact useful in measuring an important aspect of the work situation, (2) for assessment of the present state of organizaticnal functioning, or (3) for use in feeding back the data to respondents, and helping them recall specific instances in their work environment which are related to the area considered in the item.

Indices. An index is a unit of data formed by grouping several related survey items together and computing an average score for these items. The formaticn of indices makes the survey more efficient because it is often difficult to measure a domain or specific area of organizational functioning with a single survey item. Several items may be needed to adequately measure the various facets of a larger or more general domain of functioning. In order to decide which items may profitably be included in a given index, it is helpful if the items have been shown by previous data collection and research to be related. When such evidence is not available, however, one can tentatively group items with related content to see whether they do relate on the basis of the data ycu are collecting. The main concern in grouping items to form indices is that the items should not be asking the same basic question, but neither should they be so unrelated that they do not measure facets of a common domain of functioning. Ideally, items within indices should fulfill both the criteria of (1) related content and (2) relatedness tested by statistical clustering.

Dimensions. A dimension is a unit of data formed by grouping related survey items together and computing an

average score for these items. The indices described abo are sometimes grouped into dimensional areas for more efficiency in covering an area of concern. Sometimes an index is allowed to stand alone to form a dimension because no other indices have been shown to be related.

Norms. Once the survey has been administered to many people from several organizations (or commands), all of the data can be summarized together to obtain survey norms. Basically, norms may be described as average scores from all respondents on all survey items and indices. These scores are used by smaller groups of respondents - say, a single organization, command, or work group - as a standard against which their scores may be compared. When a smaller group compares its scores to those of the larger group, it gets an indication of how well it is doing relative to the norms. In order for this comparison to be meaningful, the standard group and the smaller group should have some characteristics in common: both groups should have similar organizational characteristics such as hierarchical levels or general job types; both groups should have been given the same survey under similar conditions; and both groups should have provided survey data at reasonably close time points. Most important, the standard group should be large enough to represent feelings of all people who have not been surveyed. In other words, the more people and organizations included in the norms, the more the norms provide a realistic and representative indication of their jobs as measured by the survey.

D. DEVELOPMENT OF THE NAVY'S HRM SURVEY

The HRM Survey is similar to the Survey of Organizations (SOO) (Taylor and Bowers, 1972) developed by the University

of Michigan's Institute for Social Research. The SOO items were constructed to assess various facets of organizational behaviors and were based on Likert's (1961,1967) metatheory of organizational behavior. Likert theorized that job satisfaction and performance are the result of organizational climate and leadership behaviors. The construct of organizational climate is seen as a multidimensional phenomenon and perhaps can be most clearly understood in terms of Taguiri and Litwin's (1968) definition: "Climate is a relatively enduring quality of the internal environment of an organization that (a) is experienced by its members, (b) influences their behavior, and (c) can be described in terms of the values as a particular set of characteristics (or attributes) of the organization." (p. 27)

From the standpoint of Likert's theory, organizational climate and leadership are viewed as causal variables, while job satisfaction and performance are resultant variables. Peer leadership and emergent processes (work group behavior) are theorized to be intervening variables. Taylor and Bowers (1972) and Franklin (1973, 1974) presented initial evidence supporting the postulated causal flow sequence of Likert's model. However, systematic verification of the causal hypothesis requires additional research.

While the theoretical and developmental work on the SOO was based on civilian data, it has also been administered to Navy populations as part of a study to assess the impact of changing work (life) values and preferences on Navy managerial methods. In a summary of the findings of the first 2 years of the study, Bowers and Bachman (1974) concluded that Likert's model is reasonably applicable and valid for both Navy and civilian organizations. Likewise, Franklin (1974), in assessing the causal flow model, obtained results comparable to the earlier analyses based on

civilian data. The strongest difference between the civilian and Navy samples was that peer leadership appeared to be a more critical link to group processes within the Navy. Also, Drexler and Bowers (1973) reported that organizational conditions, as measured by the SOO, accounted for significant proportions of the variance in reenlistment rates. Using ships and air squadrons as the basic units of analyses, they found positive correlations between all survey dimensions and actual reenlistment rates for those commands.

The Navy HRM Survey's structural model was the same as the SOO's. The Navy Survey was fabricated as follows: (1) many items from the SOO were adapted for use with Navy personnel through changes in terminology (e.g., organization was termed command), (2) additional items were generated from earlier efforts by Navy specialists in command development programs, and (3) items specific to contemporary social areas and programs were added as diagnostic aids for directing subsequent efforts within a command. Once the initial Navy survey questionnaire was constructed, it was modified by subsequent statistical analyses (Drexler, 1974) to yield Form X. The Navy Personnel Research and Development Center (NPRDC) subsequently added minor refinements to produce Form 09, (Bureau of Naval Personnel (Bupers) 5314-6 Publication Control Number (PCN) 09) the questionnaire used in this study. Because of the large overlap in questions between the two surveys, research findings from the SOO should be applicable to the current HRM Survey. Navy-based studies must be conducted to demonstrate the comparability of the two instruments. Crawford and Thomas (1975) predict that the considerable body of research on both the construct and predictive validity of the SOO appears to support the likelihood that similar results will be found with the Navy HRM Survey.

In Appendix A references are made to the specific survey which has been devised for use throughout Navy commands. The terms and concepts mentioned thus far will become more sharply focused as they are exemplified in the HRM Survey.

IV. THESIS HYPOTHESIS

> Personal feelings and reactions of members of the organization are probably the most useful way to find out whether the organization is bringing about the desired change. If people are reporting examples of change, then it is quite likely that the effect of the intervention is positive and useful.
>
> Glen H. Varney, 1977

It hopefully has been seen from the previous chapter that the Navy has involved itself in an intensive and comprehensive effort to effect controlled and positive organizational change. The scope of this effort is Navy wide, involving thousands of manhours per year. In view of the degree of resource allocation involved in the Navy's Organizational Development (OD) effort it is, and should be, necessary to ask the question, how well are we doing? The problem becomes, how does one answer the question. Before an evaluative answer can be formulated it is necessary to establish the criteria by which the evaluation is to be considered. The evaluative criteria may be objective, subjective or a combination of both. To draw the parallel between the evaluation of the Navy's OD efforts using survey-guided development and civilian OD efforts, we need look no farther than Rensis Likert's work with the SOO. (Likert, 1967) Over several years, his studies indicated that successful utilization of survey-guided development resulted in significant positive changes in the organization of both an objective and subjective nature.

Viewed from the performance criteria stand-point, significant research has already been conducted with

positive results. The Navy Personnel Research and Development Center was recently tasked with developing and evaluating measures of organizational effectiveness for Navy units in support of determining how the HRM cycle was impacting on unit readiness. Stated as a hypothesis, if the HRAV is effective then there should be a significant positive change in unit organizational effectiveness as measured by the HRM Survey instrument. To date at least four studies have been reported (Thomas and Crawford, 1977) which attempted to compare outcomes on selected performance criteria of units participating in the HRAV and units not participating. Although each of the four studies were conducted independently they shared a commonality of concept and statistical design. Only operational units (surface and air) were used in the studies. The basic design for the studies was to compare differences between performance measurement in a pre-HRAV and post-HRAV time frame for units having completed HRAV's and the same performance measurements in the same time frames for units not participating in HRAV's. The performance measurements selected for the studies were: 1. nonjudicial punishment rates (NJP), 2. first term reenlistment and retention rates, 3. performance in Refresher Training, and 4. readiness levels reported on NAVFORSTAT reports.

The findings of these studies are reassuring. As compared to the control units (those not participating in an HRAV) significant improvements were noted on three of the four performance measures listed above, NJP was the exception. First term reenlistment and retention rates for HRAV units increased by 7% compared to 0.2% for control units and first term retention rates rose 6.2% compared to a decline of 2.4% for control units. Although these increases were not sustained over the entire cycle, the rate of decline was lower for HRAV units than for the control units. In the Refresher Training study it was found that the

weighted averages for the HRAV units were significantly higher than the weighted averages of the control group and in fact, the experimental group had higher individual scores in five of the nine grading areas than the control group.

In the NAVFORSTAT study it was found again that the HRAV units had significantly higher ratings in Overall Readiness (R) and Equipment (E) areas than did the control units and further comparison of "before" and "after" ratings showed that 53% of the HRAV units improved compared to 25% of the control units. The conclusion drawn from these four studies seemed fairly obvious, HRAV completion by operational units had a small but significant impact (favorable) in terms of changing organizational performance. In simple terms that meant that the HRAV was doing something good and in a restricted way, what it was doing could be measured. Obviously more studies and other performance measures are needed to better define the parameters of change. Studies are in progress to do just that.

If the assumption is made that the efforts of the Human Resource Management Support System are successful then it would appear logical that it could be further assumed that subsequent survey results would reflect that success. Since the HRM survey is an attitudinal survey, it is implicit in the concept of survey guided development that attitudes are extremely important in managing organizational change. Since the survey responses are based on a five point Likert scale and are structured so that positive attitudes (good) are scored higher numerically than negative (poor) attitudes. Thus, it could be predicted that a successful HRAV would result in improved attitudes and subsequently higher unit means on future survey scores.

Historically, survey results using the SOO in civilian business organizations have borne out that prediction.

Likert's work (Likert, 1967) indicated that for a short time after the initial survey, scores would tend to drop due to the increased sensitivity of the workers to the issues addressed by the questionnaire. Scores would then increase above the initial level as expectations were met and positive change was experienced by the organization. Over time those organizations which experienced positive changes in organizational performance measures (i.e. increased profit margins, increased production output, etc.) also showed a positive trend of attitude change as reflected in higher survey scores.

Since the Navy survey instrument is very similar to the SOO and since positive changes in organizational performance measures have been observed as a result of interventions using survey guided development techniques, only a corresponding increase in aggregate survey scores over time has to be shown to draw a precise parallel between successful efforts in the civilian world and the Navy's efforts and thereby validate the concept for Navy usage.

It is the purpose of this thesis to determine, with a high degree of objective validity, whether or not a significant positive trend existed over time on the HRM survey results.

The operational hypothesis of the study is that: given the evidence of positive change as the result of HRAV efforts, a concurrent positive correlation between survey scores and time would exist. The null hypothesis was that no significant correlation existed.

The following assumptions regarding the hypothesis were made:

1. The historical HRM survey data is a statistically valid source of attitudinal change measurement over time.

2. A three year period of time is sufficient to allow relatively accurate predictions of trends.

3. Given a large sample size, the numerous modifying variables affecting scores at a unit level at particular times, will tend to cancel out over time.

Drawing on the computer data bank maintained at the Naval Postgraduate School, it was found that historical data existed on twenty (20) operational units (tape #148) covering three HRM cycle periods with a total of 11,171 cases. These units included various surface ships, air squadrons and one submarine. Also included were three overseas shore facilities. Data included both West coast and East coast units with the West coast having a slightly higher representation. All units were considered operational units and therefore the Naval Shore Establishment was not represented. All units had completed three HRAV's.

The time frame under consideration was from early 1974 thru January, 1977. Only data derived from HRM survey form 9 (Bupers 5314-6 PCN 09) were utilized. For the purpose of data analysis the time unit used was the HRM cycle. The eighteen month cycle was designated for statistical purposes as waves due to the overlap of actual HRAV completion dates. Three data points, Wave I, Wave II, and Wave III were used to represent the aggregate cycle completions.

The sample of twenty units used in this study was not selected at random and the authors freely admit the possibility of sample bias. However, even though the external validity of the findings could be questioned on the basis of sample bias, it was not considered inappropriate to attempt to generalize to the population of the operational Navy on the following basis. Comparison of the means of the

sample scores used for this study with normative data for the Navy as a whole (obtained from NPRDC, San Diego) showed the sample means to be slightly higher. This was intuitively predicted since the selection bias should have been towards the "better" units. That is to say that those units having completed the most HRAV's should have shown the greatest change and the largest absolute means.

Therefore, if in fact, the "better" units show no significant positive trend then it should be safe to make the generalization that units having completed fewer HRAV's (the rest of the the operational Navy) should show a corresponding lack of positive trend. Of course the converse generalization, that a positive trend evidenced by the sample units would tend to indicate a corresponding trend in the operational Navy overall, would not be necessarily valid at all. In fact, if the Navy overall exhibited the same positive trend as the sample then the entire concept of a causal relationship between positive attitude changes and successful efforts by the HRMSS would have to be rejected.

As indicated by the title, the statistical methodology employed in this thesis was a nonparametric approach to the testing of the alternative hypothesis. There were at least three major reasons for choosing a nonparametric approach versus the classical parametric method. The first and possibly the most important reason was that the HRM survey utilizes an ordinal scale for scoring. For that reason the distribution of scores is non-continuous and the basic assumption of the classical parametric statistical model, that the variables involved must have been measured on at least an interval scale, is not valid. It should be noted that in compiling the data for this study sample means were collected for use. It was well understood that in the case of an ordinal scale any non-integer mean was statistically

nondescriptive. The purpose in utilizing mean scores was to facilitate the assignment of ranks to the variables in question.

The second reason for rejecting the parametric approach was that during the initial data collection phase of the study it was found that the assumption of homoscedasticity was invalid. That is to say it was found that the variance for any given variable was not equal (or in a known ratio) either between units in a particular wave or between waves for the aggregate samples. In fact, the amount of change in variance between waves on several question variables exceeded the mean variance by as much as ten percent, both positively and negatively.

The third reason was more subtle but equally valid. There are suitable non-parametric statistical tests for treating samples made up of observations from several different populations, however, none of the parametric tests can handle such data without requiring the researcher to make some seemingly unrealistic assumptions. The fact that the sample used in the study was comprised of various units representing Naval air, surface, and submarine communities does not necessarily invalidate the assumption that all elements were driven from the same population. It was felt intuitively that the possibility of implicit attitudinal differences between the communities required the exercise of considerable caution in granting the validity of that assumption. Therefore, a non-parametric test was the logically cautious choice.

To further clarify the methodology it should be understood that the alternative hypothesis of this thesis did not address trends over individual variables within the HRM survey. Rather, the concern and area of analysis revolved around the survey as a composite identity or whole.

It was the explicit intention of the authors to show (or fail to show) a significant positive trend for the survey overall irrespective of the direction or magnitude of possible trends of the internal component variables (questions, indices, dimensions).

The prediction that a positive overall trend would result as the efforts of an HRAV being institutionalized over time was based on the consideration that no effort in a specific area would result in a negative impact on another area. That is to say, efforts to improve communications, for example, would not adversely effect command climate. It was reasoned therefore that over time the cumulative impact should have been positive for those areas towards which an HRM effort was directed. Areas not addressed during the HRAVs should, at worst, have remained constant but in no case should the cumulative trend have been negative. It should be reemphasized that the assumption was made that the effects of moderating external variables could be considered constant due to the large sample size and the dampening effects inherent in the relatively long time frames under consideration.

For the interested reader, appendix B contains a direct comparison of equivalent parametric and nonparametric tests (Pearson versus Spearman). The correlation coefficients given are for individual survey questions with wave. Notice should be given to the fact that although in neither case were the coefficients large, the difference between coefficients on the same questions are in many cases proportional.

V. METHODOLOGY

Before any statistical testing could be accomplished the authors felt it was necessary to develop a logical series of descriptive statistics in an attempt to visually display the data. Due to the manner in which the data was compiled and the type of analysis under consideration, a matrix format was selected for tabular continuity and ease of visual inspection. As stated previously, arithmetic means were tabulated as a first step to organize the data in a manner amenable to conversion to rank ordering. Although it was recognized that means based on an ordinal scale are not truly descriptive, the gross change in means between waves was felt to be adequate for an initial estimate of the amount and direction of change to be expected. Fig 3 illustrates the array of question means for each wave. Three decimal places were selected to distinguish small amounts of change for later significance testing.

Question	Wave I	Wave II	Wave III
1	2.795	2.867	2.866
2	3.207	3.289	3.318
3	3.116	3.064	3.094
4	2.983	2.938	2.916
5	2.860	2.860	2.885
6	2.707	2.483	2.404
7	3.429	3.315	3.322
8	3.017	3.020	3.074
9	2.770	2.726	2.806
10	2.977	2.736	2.754
11	2.685	2.758	2.790
12	3.093	3.110	3.200
13	2.788	2.743	2.715
14	2.857	2.775	2.778
15	2.425	2.418	2.392
16	3.842	3.913	3.964
17	3.737	3.706	3.764
18	3.685	3.714	3.747
19	3.433	3.507	3.535
20	3.440	3.440	3.454
21	3.199	3.221	3.231
22	3.714	3.647	3.645
23	3.727	3.721	3.754
24	3.158	3.210	3.275
25	3.063	3.070	3.117
26	3.159	3.240	3.286
27	4.074	4.010	4.058
28	3.684	3.665	3.699
29	3.409	3.407	3.439
30	3.232	3.169	3.243
31	2.976	2.956	3.010
32	3.255	3.084	3.105
33	3.537	3.449	3.504
34	2.951	3.024	3.122
35	2.914	2.920	2.953
36	3.192	3.202	3.277
37	3.292	3.272	3.304
38	3.430	3.503	3.560
39	3.122	3.145	3.229
40	3.590	3.509	3.617
41	3.356	3.358	3.447
42	3.471	3.428	3.471
43	3.405	3.457	3.429
44	3.798	3.776	3.770

Figure 3 - SURVEY QUESTION STATISTICAL MEANS

Question	Wave I	Wave II	Wave III
45	3.905	3.932	3.962
46	3.447	3.427	3.514
47	3.532	3.460	3.572
48	3.836	3.799	3.870
49	3.692	3.793	3.775
50	3.569	3.504	3.506
51	3.155	3.048	2.992
52	3.537	3.523	3.568
53	3.144	3.124	3.234
54	3.299	3.198	3.195
55	3.292	3.263	3.280
56	2.792	2.719	2.724
57	3.633	3.504	3.521
58	3.190	2.923	2.966
59	3.012	2.962	3.039
60	3.278	3.328	3.349
61	3.464	3.288	3.309
62	2.621	2.608	2.620
63	3.212	3.099	3.129
64	3.649	3.506	3.573
65	3.252	3.182	3.211
66	3.396	3.200	3.203
67	3.497	3.440	3.474
68	3.440	3.165	3.151
69	3.369	3.196	3.291
70	3.194	3.113	3.133
71	3.184	2.995	3.064
72	3.260	2.966	2.971
73	2.896	2.641	2.671
74	3.740	3.690	3.768
75	3.419	3.258	3.336
76	3.574	3.377	3.412
77	3.513	3.534	3.508
78	3.208	3.086	3.223
79	2.609	2.573	2.635
80	2.671	2.718	2.873
81	3.640	3.648	3.717
82	3.374	3.201	3.260
83	3.797	3.594	3.612
84	2.944	3.111	3.117
85	3.584	3.437	3.468
86	3.013	2.935	2.902
87	3.924	3.881	3.928
88	3.531	3.894	3.895

Figure 4 - SURVEY QUESTION STATISTICAL MEANS CONTINUED

Fig 5 and Fig 6 display the matrices of the means on the index and dimension variables respectively. These means were computed by averaging the component index means for the dimensions. Again the means were carried to three decimal places to distinguish small differences and to facilitate rank ordering.

Index	Wave I	Wave II	Wave III
1	3.039	3.073	3.093
2	2.850	2.760	2.735
3	3.072	3.020	3.067
4	2.886	2.837	2.865
5	2.641	2.596	2.585
6	3.674	3.710	3.753
7	3.320	3.331	3.343
8	3.721	3.684	3.670
9	3.127	3.173	3.226
10	3.722	3.694	3.732
11	3.250	3.165	3.216
12	2.933	2.972	3.038
13	3.305	3.326	3.380
14	3.385	3.360	3.441
15	3.703	3.722	3.630
16	3.490	3.444	3.543
17	3.462	3.427	3.449
18	3.042	2.991	3.002
19	3.278	3.130	3.175
20	3.121	3.075	3.093
21	3.304	3.137	3.170
22	3.578	3.442	3.505
23	3.128	3.112	3.191
24	3.452	3.436	3.455

INDICES

1. Communications flow
2. Decision making
3. Motivation
4. Human resources emphasis
5. Lower level influence
6. Supervisory support
7. Sup. teamwork
8. Sup. goal emphasis
9. Sup. work facilitation
10. Peer support
11. Peer teamwork
12. Peer work facilitation
13. Peer problem solving
14. Work group coordination
15. Work group readiness
16. Work group discipline
17. Satisfaction
18. Integration of men and mission
19. Training
20. General
21. Equal opportunity/Race relations
22. Drug abuse
23. Alcoholism prevention
24. Community interrelationships

Figure 5 - STATISTICAL MEANS OF THE INDICES

Dimension	Wave I	Wave II	Wave III
1	2.898	2.857	2.869
2	3.461	3.475	3.498
3	3.303	3.289	3.342
4	3.526	3.509	3.538
5	3.462	3.427	3.449
6	3.042	2.991	3.002
7	3.278	3.130	3.175
8	3.121	3.075	3.093
9	3.304	3.137	3.170
10	3.578	3.442	3.505
11	3.128	3.112	3.191
12	3.452	3.436	3.455

DIMENSIONS

1. Command climate
2. Supervisory leadership
3. Peer leadership
4. Work group process
5. Satisfaction
6. Integration of men and mission
7. Training
8. General
9. Equal opportunity/Race relations
10. Drug abuse
11. Alcoholism prevention
12. Community interrelationships

Figure 6 - STATISTICAL MEANS OF THE DIMENSIONS

A. MATCHED-PAIRS ANALYSIS

Visually inspecting the direction and magnitude of change of the arithmetic means yielded little insight into significant trends due to small changes in absolute value and the fact that the direction of changes did not appear to be consistent. It was of course possible to compute a rough arithmetic approximation of overall change by simply averaging all scores for each wave and inspecting the degree of change between these values. Wave I yielded an average score of 3.333, Wave II yielded an average score of 3.299, and Wave III yielded an average score of 3.330. Using these values as an approximation of the magnitude and direction of overall change in survey scores, it appeared that a slight negative trend existed between Wave I and Wave II and then an equal but positive trend between Wave I and Wave III returning the level overall to the Wave I value. The net result overall would seem to be no significant change.

If the assumption is made that observations on the same variable between any two Waves are derived from related samples then the Sign test (Siegal, 1956) can be utilized to test for significance of change in the direction of scores. The null hypothesis tested by the sign test is that $p(Sw1>Sw2) = p(Sw1<Sw2) = 0.5$ where Sw1 is the score on Wave I and Sw2 is the score on Wave II for the same variable. That is, Sw1 and Sw2 are the two scores for a matched pair. Another way of stating the null hypothesis is: the median difference is zero.

In applying the sign test, the focus is on the direction of the differences between all of the matched pairs in any two Waves. Under the null hypothesis it would be expected

that the number of pairs in which the difference is positive would equal the number of pairs in which the difference is negative. The null hypothesis would be rejected if too few differences of one sign occurred. The probability associated with the occurance of a particular number of positives and negatives can be determined by reference to the binomial distribution with P = Q = 0.5, where N = the number of pairs. If a matched pair shows no difference it is dropped from the analysis and N is thereby reduced.

The power-efficiency of the sign test is about 95% for N = 6, but declines as the sample size increases to an eventual (asymptotic) efficiency of about 63%. Due to the relationship between the size of N and the power-efficiency of the test, no attempt was made to use the matrix of question means. Instead the test was applied to the means of dimensions thereby favoring greater power efficiency. Fig 7 shows the sign distribution between Wave I and Wave II for dimension means.

Dimension	Wave I Mean	Wave II Mean	Sign of difference
1	2.898	2.857	-
2	3.461	3.475	+
3	3.303	3.289	-
4	3.526	3.509	-
5	3.462	3.427	-
6	3.042	2.991	-
7	3.278	3.130	-
8	3.121	3.075	-
9	3.304	3.137	-
10	3.578	3.442	-
11	3.128	3.112	-
12	3.452	3.436	-

Figure 7 - WAVE I AND WAVE II MATCHED DIMENSIONS AND CHANGE SIGN DISTRIBUTION

The sum of the positive signs is equal to one (1) and the sum of the negative signs is equal to eleven (11). The aggregate survey means for each Wave indicated a drop in the overall value for Wave II compared to Wave I. Therefore, a one tailed test was utilized based on a predicted negative trend. Setting the level of significance for the test at alpha equal to 0.05 it was found using the appropriate tables (Siegel, p. 250 table D) that the probability of a distribution of one positive and eleven negatives or an even more extreme case under the null hypothesis was $P = 0.003$ (one-tailed). Therefore the null hypothesis was rejected at the 0.05 level of significance.

Fig 8 shows the sign distribution between Wave II and Wave III for dimension means.

Dimension	Wave II Mean	Wave III Mean	Sign of difference
1	2.857	2.869	+
2	3.475	3.498	+
3	3.289	3.342	+
4	3.509	3.538	+
5	3.427	3.449	+
6	2.991	3.002	+
7	3.130	3.175	+
8	3.075	3.093	+
9	3.137	3.170	+
10	3.442	3.505	+
11	3.112	3.191	+
12	3.436	3.455	+

Figure 8 - WAVE II AND WAVE III MATCHED DIMENSIONS AND CHANGE SIGN DISTRIBUTION

As before the direction of change was predicted (in this case positive) therefore a one-tailed test of significance was utilized. The sum of positives was equal to twelve and the sum of the negative signs was equal to zero. No tie matches occurred therefore N remains equal to 12. Again using alpha equal to 0.05 and entering the appropriate tables, (p. 250 table D) the probability under the null hypothesis of this distribution was found to be P = 0.001. Since P = 0.001 is less than 0.05, the null hypothesis can be rejected and the conclusion drawn that a significant positive change occurred between Wave II and Wave III.

Fig 9 shows the sign distribution between Wave I and Wave III for dimension means.

Dimension	Wave I Mean	Wave III Mean	Sign of difference
1	2.898	2.869	-
2	3.461	3.498	+
3	3.303	3.342	+
4	3.526	3.538	+
5	3.462	3.449	-
6	3.042	3.002	-
7	3.278	3.175	-
8	3.121	3.093	-
9	3.304	3.170	-
10	3.578	3.505	-
11	3.128	3.191	+
12	3.452	3.455	+

Figure 9 - WAVE I AND WAVE III MATCHED DIMENSIONS AND CHANGE SIGN DISTRIBUTION

The prediction in this case was that no significant change had occurred between Wave I and Wave III. Therefore failure to reject the null hypothesis was anticipated. The sum of the positive signs was equal to 5 and the sum of the negative signs was equal to 7, with no ties N = 12. From the probability tables it was found that under the null hypothesis the probability of this distribution (using a two-tailed test) was p = 0.774. As before, the level of significance being tested for was alpha equal to 0.05 and since p = 0.774 is greater than 0.05 the result was failure to reject the null hypothesis as predicted.

Momentarily disregarding the level of significance used for the testing, it should be noted that the general direction of change indicated by the sign tests was in agreement with the change predicted by the survey arithmetic means. That is, the overall trend between Wave I and Wave II was negative in direction and the trend (overall) between Wave II and Wave III was positive. Although the net change between Wave I and Wave III was not significant the sign test has revealed the possibility of a curvilinear relationship which may have been beginning to trend upward by the end of Wave III.

The test just used, the sign test, utilized information simply about the direction of the differences within pairs. If the relative magnitude as well as the direction is considered, a more powerful test can be made. The Wilcoxon matched-pairs signed-ranks test gives more weight to a pair which shows a large difference between the two waves under consideration than to a pair which shows a small difference. To proceed, let di = the difference score for any matched pair. Each pair has one di. After all di's are determined they are then ranked without regard to sign with the smallest di ranking 1, the next smallest 2, etc. Then to each rank the sign of the difference is affixed. If any two

Waves are equivalent, that is, if the null hypothesis is true, then the sum of the positive ranks and the sum of the negative ranks should be approximately equal. Said another way,

$$H_0 : \Sigma\text{-di} \simeq \Sigma\text{+di}$$

If the sums of the ranks are very different it can be inferred that the Waves differ from each other at some level of significance. The null hypothesis can be rejected if the sum of either the positive or negative di's is too small at the chosen level of significance.

The power-efficiency of this test is near to 95% for N less than or equal to 25 therefore it can be applied to the means of survey indices. Since the test statistic T, the smaller sum of like-signed ranks, has its probabilities computed up to values of N = 25 it will not be necessary to convert T to a z score, thus no assumption of normality is necessary for the test.

As in the sign test, the level of significance is set at alpha equal to 0.05. In the event of tied ranks the average of the ranks which would have been assigned was substituted. T was computed from Wave I and Wave II, Wave II and Wave III, and Wave I and Wave III. Fig 10 shows the computation of T for Wave I and Wave II.

Index	Wave I	Wave II	d	Rank of d	Rank with less frequent sign
1	3.039	3.073	0.034	8	8
2	2.850	2.760	-0.090	-21	
3	3.072	3.020	-0.052	-19	
4	2.886	2.837	-0.049	-17	
5	2.641	2.596	-0.045	-13	
6	3.674	3.710	0.036	10	10
7	3.320	3.331	0.011	1	1
8	3.721	3.684	-0.037	-11	
9	3.127	3.173	0.046	15	15
10	3.722	3.694	-0.028	-7	
11	3.250	3.165	-0.085	-20	
12	2.933	2.972	0.039	12	12
13	3.305	3.326	0.021	5	5
14	3.385	3.360	-0.025	-6	
15	3.703	3.722	0.019	4	4
16	3.490	3.444	-0.046	-15	
17	3.462	3.427	-0.035	-9	
18	3.042	2.991	-0.051	-18	
19	3.278	3.130	-0.148	-23	
20	3.121	3.075	-0.046	-15	
21	3.304	3.137	-0.167	-24	
22	3.578	3.442	-0.136	-22	
23	3.128	3.112	-0.016	-2.5	
24	3.452	3.436	-0.016	-2.5	
					T = 55

Figure 10 - WAVE I TO WAVE II WILCOXON MATCHED-PAIR SIGNED-RANK TEST

Summing the ranks with positive signs a value of T = 55 is computed. Entering the appropriate tables (Siegel, p. 254, table G) it is found that the probability under the null hypothesis of this value of T (one-tailed test) is $p < 0.005$, therefore the test rejects the null hypothesis at the 0.05 level of significance.

Fig 11 shows the computation of T for Wave II and Wave III.

Index	Wave II	Wave III	d	Rank of d	Rank with less frequent sign
1	3.073	3.093	0.020	7	
2	2.760	2.735	-0.025	-9	9
3	3.020	3.067	0.047	15	
4	2.837	2.865	0.028	10	
5	2.596	2.585	-0.011	-2	2
6	3.710	3.753	0.043	13	
7	3.331	3.343	0.011	2	
8	3.684	3.670	-0.014	-4	4
9	3.173	3.226	0.053	17	
10	3.694	3.732	0.038	12	
11	3.165	3.216	0.051	16	
12	2.972	3.038	0.066	20	
13	3.326	3.380	0.054	18	
14	3.360	3.441	0.081	22	
15	3.722	3.630	-0.092	-23	23
16	3.444	3.543	0.099	24	
17	3.427	3.449	0.022	8	
18	2.991	3.002	0.011	2	
19	3.130	3.175	0.045	14	
20	3.075	3.093	0.018	5	
21	3.137	3.170	0.033	11	
22	3.442	3.505	0.063	19	
23	3.112	3.191	0.079	21	
24	3.436	3.455	0.019	6	
					$\overline{T} = 38$

Figure 11 - WAVE II TO WAVE III WILCOXON MATCHED-PAIR SIGNED-RANK TEST

Summing the negative ranks a value of T = 38 is computed with N = 25. Entering the tables to find the probability of T associated with a one-tailed test with alpha equal to 0.05 it was found to be $p < 0.005$. Since $0.005 < 0.05$ the null hypothesis was rejected at the 0.05 level. Therefore a significant positive trend of change could be inferred between Wave II and Wave III.

Fig 12 shows the computation of T for Wave I and Wave III.

Index	Wave I	Wave III	d	Rank of d	Rank with less frequent sign
1	3.039	3.093	0.054	12	12
2	2.850	2.735	-0.005	-23	
3	3.072	3.067	-0.005	-2	
4	2.886	2.865	-0.021	-5	
5	2.641	2.585	-0.056	-13.5	
6	3.674	3.753	0.079	19	19
7	3.320	3.343	0.023	6	6
8	3.721	3.670	-0.051	-10	
9	3.127	3.226	0.099	20	20
10	3.722	3.732	0.010	3	3
11	3.250	3.216	-0.034	-8	
12	2.933	3.038	0.105	21	21
13	3.305	3.380	0.075	18	18
14	3.385	3.441	0.056	13.5	13.5
15	3.703	3.630	-0.073	-16.5	
16	3.490	3.543	0.053	11	11
17	3.462	3.449	-0.013	-4	
18	3.042	3.002	-0.040	-9	
19	3.278	3.175	-0.103	-22	
20	3.121	3.093	-0.028	-7	
21	3.304	3.170	-0.134	-24	
22	3.578	3.505	-0.073	-16.5	
23	3.128	3.191	0.063	15	15
24	3.452	3.455	0.003	1	1
					T = 139.5

Figure 12 - WAVE I TO WAVE III WILCOXON MATCHED-PAIR SIGNED-RANK TEST

Summing the negative ranks a value of T = 139.5 is computed. Entering the tables with N = 24 it was found that the critical value of T for alpha equal to 0.05 was T = 81. Since 139.5 > 81 this test failed to reject the null hypothesis at the 0.05 level.

It was of considerable interest to note that the Wilcoxon test yielded precisely the same results in this case as the less powerful sign test. From a descriptive point of view the Wilcoxon test yields more insight. The utilization of the Wilcoxon matched-pair signed-rank test clearly showed the decisiveness of the positive change between Wave II and Wave III as well as the fact that the lack of significant overall change between Wave I and Wave III was the result of the magnitude of the negative components rather than the number of negative ranks. It should be clarified that the two tests used to this point, the sign test and the Wilcoxon matched-pairs signed-ranks test, were tests of two related samples. The concept of relatedness in this sense was that the groups represented by the three Waves served as their own controls. The implication in related samples of this type is the statistical assumption that each Wave represented the same group taking the survey at three different times. This assumption is perfectly valid and implicitly necessary when the consideration is given to the fact that the units from which the data was collected were the same in each Wave. Although the Waves are related in a statistical sense, it would have been unrealistic to ignore the fact that the individual constituency of the various units was constantly changing over the period encompassed by the study. In fact, given an average attrition rate of 25% per annum for the Navy, in three years it could be expected that the crew of any given unit could have changed by approximately 75%.

Given this consideration, a strong argument could be made for the fact that the three Waves actually represent three independent samples drawn from the same population. The design assumption under this argument would be that samples (Waves) arose from the random treatment of members of some sample whose origins are arbitrary. It was felt by the authors that the argument for related samples was stronger but that for the sake of considering the alternate possibility tests for independent samples would be included in the analysis.

The Kolmogorov-Smirnov two-sample test is a test of whether two independent samples have been drawn from the same population (or from populations with the same distribution). The one-tailed test can be used to decide, at a given level of significance, whether or not the values of the population from which one sample was drawn are stochastically larger than the values of the population from which the other sample was drawn. That is, it can be used to test the prediction that the scores of one group will be "better" than those of the other group. This test is concerned with the agreement between two cumulative distributions, e.g., the agreement between two sets of sample values.

When compared with the t test, the Kolmogorov-Smirnov test has high power-efficiency (about 96%) for small sample sizes (N is less than or equal to 40). This test seems to be more powerful in all cases than either the chi-square test or the median test, hence its selection for this analysis.

To apply the test a cumulative frequency distribution was constructed for each Wave under consideration using the same intervals for each distribution (in this case the selected interval was 0.05). The interval selection in this

test case was arbitrary but was felt to be sufficiently small (that is, there was a sufficient number of intervals) that the maximum vertical deviation of the two cumulative step functions would not be obscured. The test statistic, D, can be defined as:

$$D = \text{maximum}\{SN_1(X) - SN_2(X)\}$$

where

$$SN(X) = K/N_1$$

K is equal to the number of scores equal to or less than X and N1 is equal to the number of scores in the sample. Therefore the cumulative step function of the other Wave in the same interval (X) yields D when that difference is the largest of all the differences. For the one-tailed test this difference must be the largest in the predicted direction.

The means of survey indices were selected for this test to include the largest number of scores without exceeding N = 40. In this case N1 = 24 and N2 = 24 therefore N1 = N2 and both are less than 40 which satisfied the requirements necessary tc use the tabulated values of D for significance testing. Fig 13 shows the calculation of D for Wave I and Wave II. Based on the results of the previous tests, the prediction was that Wave II scores would be less than Wave I scores. The largest negative difference then is the value of D.

Interval	S(X) Wave I	S(X) Wave II	$S(X)_I - S(X)_{II}$
2.551-2.600	0	1/24	-1/24
2.601-2.650	1/24	1/24	0
2.651-2.700	1/24	1/24	0
2.701-2.750	1/24	1/24	0
2.751-2.800	1/24	2/24	-1/24
2.801-2.850	2/24	3/24	-1/24
2.851-2.900	3/24	3/24	0
2.901-2.950	4/24	3/24	1/24
2.951-3.000	4/24	5/24	-1/24
3.001-3.050	6/24	6/24	0
3.051-3.100	7/24	8/24	-1/24
3.101-3.150	10/24	11/24	-1/24
3.151-3.200	10/24	13/24	-3/24
3.201-3.250	11/24	13/24	-2/24
3.251-3.300	12/24	13/24	-1/24
3.301-3.350	15/24	15/24	0
3.351-3.400	16/24	16/24	0
3.401-3.450	16/24	20/24	-4/24
3.451-3.500	19/24	20/24	-1/24
3.501-3.550	19/24	20/24	-1/24
3.551-3.600	20/24	20/24	0
3.601-3.650	20/24	20/24	0
3.651-3.700	21/24	22/24	-1/24
3.701-3.750	24/24	24/24	0

Figure 13 - KOLMOGOROV-SMIRNOV TEST FOR WAVE I AND WAVE II

As seen, K_D= 4 where K_D is the numerator of the largest difference. Comparing the critical value of $K_D$at alpha equal to 0.05 (Siegel, p. 278, table L) the critical value was found to be K_D= 9 for N = 24. Thus the test failed to reject the null hypothesis at the 0.05 level of significance resulting in the inference that the scores of Wave II are not significantly smaller than the scores of Wave I.

Fig 14 shows the calculation of D for Wave II amd Wave III. The prediction in this case was that the Wave III scores would be significantly larger than the Wave II scores, therefore the maximum D would be the D with the largest positive value.

Interval	S(X) Wave II	S(X) Wave III	$S(X)_{II} - S(X)_{III}$
2.551-2.600	1/24	1/24	0
2.601-2.650	1/24	1/24	0
2.651-2.700	1/24	1/24	0
2.701-2.750	1/24	2/24	-1/24
2.751-2.800	2/24	3/24	-1/24
2.801-2.850	3/24	3/24	0
2.851-2.900	3/24	3/24	0
2.901-2.950	3/24	3/24	0
2.951-3.000	5/24	3/24	2/24
3.001-3.050	6/24	5/24	1/24
3.051-3.100	8/24	8/24	0
3.101-3.150	11/24	8/24	3/24
3.151-3.200	13/24	11/24	2/24
3.201-3.250	13/24	13/24	0
3.251-3.300	13/24	13/24	0
3.301-3.350	15/24	14/24	1/24
3.351-3.400	16/24	15/24	1/24
3.401-3.450	20/24	17/24	3/24
3.451-3.500	20/24	18/24	2/24
3.501-3.550	20/24	20/24	0
3.551-3.600	20/24	20/24	0
3.601-3.650	20/24	21/24	-1/24
3.651-3.700	22/24	22/24	0
3.701-3.750	24/24	24/24	0

Figure 14 - KOLMOGOROV-SMIRNOV TEST FOR WAVE II AND WAVE III

The critical value of K_D is 9 (the same as before). Since $K_D = 3 < 9$ the null hypothesis failed to be rejected again at the 0.05 level of significance. The inference being that survey scores on Wave III were not significantly higher than those of Wave II.

Fig 15 illustrates the calculation of D for Wave I and Wave III. The prediction in this case was that the scores for Wave III would be higher than those of Wave I.

Interval	S(X) Wave I	S(X) Wave III	$S(X)_I - S(X)_{III}$
2.551-2.600	0	1/24	-1/24
2.601-2.650	1/24	1/24	0
2.651-2.700	1/24	1/24	0
2.701-2.750	1/24	2/24	-1/24
2.751-2.800	1/24	3/24	-2/24
2.801-2.850	2/24	3/24	-1/24
2.851-2.900	3/24	3/24	0
2.901-2.950	4/24	3/24	1/24
2.951-3.000	4/24	3/24	1/24
3.001-3.050	6/24	5/24	1/24
3.051-3.100	7/24	8/24	-1/24
3.101-3.150	10/24	8/24	2/24
3.151-3.200	10/24	11/24	-1/24
3.201-3.250	11/24	13/24	-2/24
3.251-3.300	12/24	13/24	-1/24
3.301-3.350	15/24	14/24	1/24
3.351-3.400	16/24	15/24	1/24
3.401-3.450	16/24	17/24	-1/24
3.451-3.500	19/24	18/24	1/24
3.501-3.550	19/24	20/24	-1/24
3/551-3.600	20/24	20/24	0
3.601-3.650	20/24	21/24	-1/24
3.651-3.700	21/24	22/24	-1/24
3.701-3.750	24/24	24/24	0

Figure 15 - KOLMOGROV-SMIRNOV TEST FOR WAVE I AND WAVE III

K_D = 2 in this case. Since 2 < 10 (the critical value of K_D at alpha equal to a 0.05 level of significance) the test fails to reject the null hypothesis and the inference was drawn that no significant difference exists between the scores on Wave I and those of Wave III.

B. THREE SAMPLE SIMULTANEOUS ANALYSIS

Although each of the tests conducted have been relevant to the question of significant difference between Waves, it must be realized that a serious problem exists with those types of tests. Matched-pair tests of any type are accurate only when the matched-pair change is not dependent upon an intervening variable. In effect the results of matched-pair difference tests between Wave I and Wave II and between Wave II and Wave III can be considered valid at the level of significance used, however the tests between Wave I and Wave III exhibit a compound error in that no consideration is given the effect of Wave II. The results of such tests must be considered in error to the extent that Wave II effects the total change. An example of this error effect can be illustrated by plane geometry. Take three points A, B, and C which define a straight line. The distance between points A and C is exactly equal to the sum of the distances between A and B and B and C. If, instead of a line, the three points defined a triangle with the distance of line AC being the base then the sum of lengths of AB and BC would be considerably greater than AC. How much greater would be a function of the angular displacement of B relative to AC.

The results of the sign test and the Wilcoxon matched-pairs signed-ranks test indicated an analagous triangular relationship between the scores of the three

Waves. Since it appeared that the scores on Wave II would not fall on a straight line connecting the scores of Wave I and Wave III (on the average) it was felt that the change occurring only between Wave I and Wave III was not at all truly indicative of the amount of total change. The next step was to determine whether or not general agreement existed overall between the Waves as to the relative ranking of each variable within the Wave. If in fact it was found that each Wave, treated as an independent judge, ranked the variables in the same relative order then it could be concluded that the overall change in scores was a function of trend in general and not a function of opposing trends tending to have a cancelling effect on each other. Using the geometric analogy to illustrate again, if the score on a particular variable in each Wave was treated as three points of a triangle and if overall the point defined by Wave II was below the line defined by Wave I and Wave III then superimposing all of the resulting triangles would result in a triangle describing the average shape of all the possible ones. Such a result would lead to the inescapable conclusion that the difference between Wave I and Wave III as a pair was not indicative of the total difference as represented by the sum of the changes between Wave I and Wave II and Wave II and Wave III.

Within Wave rankings on all three variable sets, questions, indices, and dimensions, were constructed as descriptive statistics. Ranks were labeled as measures of concern on the assumption that a low score on a particular variable indicated a greater degree of concern by the individual respondent while a relatively high score indicated less concern (or more satisfaction). Fig 16 illustrates the within Wave ranking of questions (appendix A lists all survey questions for comparison purposes, only question numbers are used in the tables). Ties are indicated by paranthetical markings.

	Rank	Question # Wave I	Question # Wave II	Question # Wave III
	1	15	15	15
	2	79	6	6
	3	62	79	62
most	4	80	62	79
concern	5	11	73	73
(lowest	6	6	80	13
raw	7	9	56	56
score)	8	13	9	10
	9	56	10	14
	10	1	13	11
	11	14	11	9
	12	5	14	1
	13	73	5	80
	14	35	1	5
	15	84	35	86
	16	34	58	4
	17	31	86	35
	18	10	4	58
	19	4	31	72
	20	59	59	51
	21	86	72	31
	22	8	71	59
	23	25	8	71
	24	12	34	8
	25	3	51	3
	26	39	3	32
	27	53	25	25(84)
	28	51	32	84(25)
	29	24	78	34
	30	26	63	63
	31	71	12	70
	32	58	84	68
	33	36	70	54
	34	70	53	12
	35	21	39	66
	36	2	68	65
	37	78	30	78
	38	63	65	39
	39	30	69	21
	40	65	54	53
	41	32	66	30
	42	72	82	82
	43	60	36	24
	44	55(37)	24	36

Figure 16 - HRM SURVEY QUESTIONS RANKED BY WAVE

Rank	Question # Wave I	Question # Wave II	Question # Wave III
45	37(55)	21	55
46	54	26	26
47	41	75	69
48	69	55	37
49	82	37	61
50	66	61	2
51	43	2	7
52	29	7	75
53	75	60	60
54	7	41	76
55	38	76	43
56	19	29	29
57	20(68)	46	41
58	68(20)	46	41
59	46	85	85
60	61	67(20)	42
61	42	20(67)	67
62	67	33	33
63	77	43	50
64	88	47	77
65	47	38	46
66	52(33)	50(57)	57
67	33(52)	57(50)	19
68	50	64	38
69	76	19	52
70	85	40	47
71	40	52	64
72	57	77	83
73	81	83	40
74	64	22	22
75	28	81	28
76	18	28	81
77	49	74	18
78	22	17	28
79	23	18	17
80	17	23	74
81	74	44	44
82	83	49	49
83	44	48	48
84	48	87	88
85	16	88	87
86	45	16	45
87	87	45	16
88	27	27	27

least concern (highest raw score)

Figure 17 - HRM SURVEY QUESTIONS RANKED BY WAVE CONTINUED

Fig 18 illustrates the within wave rankings of survey indices and lists the respective indices by number. As before, ties are indicated parenthetically.

	Rank	Index # Wave I	Index # Wave II	Index # Wave III
	1	5	5	5
most	2	2	2	2
concern	3	4	4	4
(lowest	4	12	12	18
raw	5	1	18	12
score)	6	18	3	3
	7	3	1	1(20)
	8	20	20	20(1)
	9	9	23	21
	10	23	19	19
	11	11	21	23
	12	19	11	11
	13	21	9	9
	14	13	13	7
	15	7	7	13
	16	14	14	14
	17	24	17	17
	18	17	24	24
	19	16	22	22
	20	22	16	16
least	21	6	8	15
concern	22	15	10	8
(highest	23	8	6	10
raw	24	10	15	6
score)				

Figure 18 - HRM SURVEY INDICES RANKED BY WAVE

Fig 19 illustrates the within wave ranking of survey dimensions with ranking of variables by number.

	Rank	Wave I	Wave II	Wave III
most	1	1	1	1
concern	2	6	6	6
(lowest raw	3	8	8	8
score)	4	11	11	9
	5	7	7	7
	6	3	9	11
	7	9	3	3
	8	12	5	5
least	9	2	12	12
concern	10	5	10	2
(highest raw	11	4	2	10
score)	12	10	4	4

Figure 19 - HRM SURVEY DIMENSIONS RANKED BY WAVE

Visual examination of the within Wave rankings tended (with some obvious exceptions) to confirm the hypothesis that in general the overall rankings remained relatively consistent from Wave to Wave. Kendall's coefficient of concordance, W, is a powerful measurement of the relation among several rankings of N groups of variables. It would have been equally effective to compute the Spearman rank correlation coefficient between each wave and then take the average of the coefficients. The computation of W is much simpler, and W bears a linear relation to the average r's taken over all groups. Thus by using W not only are all of the Waves considered simultaneously but by using the linear relationship formula

$$r_s(\text{avg.}) = \frac{KW-1}{k-1}$$

the average correlation between individual Waves can be calculated. Using the k times N tables already constructed (figures 16 and 17) for survey indices and dimensions, W was computed in the following manner.

The sum of the ranks, Rj, in each column (Wave) was found. All the Rj's are then summed and divided by N to determine the mean value of Rj. Each Rj may then be expressed as a deviation from the mean value. Then the sum of the squares of the deviations are found. The formula for W is :

$$W = \frac{S}{1/12\ K^2(N^3-N)}$$

where

s = the sum of squares of the observed deviations from the mean of Rj

k = number of sets of rankings (in this case k = 3)

N = the number of entries ranked (for the indices N = 24, for the dimensions N = 12)

and $1/12\ K^2(N^3-N)$ is equal to the maximum possible sum

of the squared deviations.

Figures 20 and 21 show the rank distribution of the index and dimension means respectively. For ease of reading the tables are presented in a N times k format, therefore the values of Rj and $(R_j-\Sigma R_j/N)^2$ appear to be for rows instead of cclumns. The values presented in the figures actually represent the sums of columns and the formula values are appropriate for the data.

Index	Wave I	Wave II	Wave III	R_j	$(R_j - \Sigma R_j/N)^2$
1	5	7	8	20	306.25
2	2	2	2	6	992.25
3	7	6	6	19	342.25
4	3	3	3	9	812.25
5	1	1	1	3	1,190.25
6	21	23	24	68	930.25
7	15	15	14	44	42.25
8	23	21	22	66	812.25
9	9	13	13	35	6.25
10	24	22	23	69	992.25
11	11	12	12	35	6.25
12	4	4	5	13	600.25
13	14	14	15	43	30.25
14	16	16	16	48	110.25
15	22	24	21	67	870.25
16	19	20	20	59	462.25
17	18	17	17	52	210.25
18	6	5	4	15	506.25
19	12	10	10	32	30.25
20	8	8	7	23	210.25
21	13	11	9	33	20.25
22	20	19	19	58	420.25
23	10	9	11	30	56.25
24	17	18	18	53	240.25

Figure 20 - RANK DISTRIBUTION OF INDEX MEANS

Dimension	Wave I	Wave II	Wave III	R_j	$(R_j - R_j/N)^2$
1	1	1	1	3	272.25
2	9	11	10	30	110.25
3	6	7	7	20	0.25
4	11	12	12	35	240.25
5	10	8	8	26	42.25
6	2	2	2	6	182.25
7	5	5	5	15	20.25
8	3	3	3	9	110.25
9	7	6	4	17	6.25
10	12	10	11	33	182.25
11	4	4	6	14	30.25
12	8	9	9	26	42.25

Figure 21 - RANK DISTRIBUTION OF DIMENSION MEANS

Solving for the dimension variables first, the sum of Rj = 244 therefore $\Sigma R_j/N = 244/12 = 20.333$ since $S=\Sigma(R_j-\frac{\Sigma R_j}{N})^2$

$$\text{then } S=\Sigma R_j-20.333 = 1239$$

Substituting in the formula,

$$W = \frac{1239}{1/12(3^2)(12^3-12)}$$

$$W = 0.9627$$

when N > 7, the formula

$$\chi^2 = K(N-1)W$$

is approximately distributed as chi square with df = N - 1. That is, the probability associated with the occurrance under the null hypothesis of any value as large as an observed W may be determined by finding chi square by the formula above and then using the chi square tables to determine the probability of chi square. The null hypothesis is that no correlation exists over the k rankings. Substituting the computed value of W into the formula the following value of chi square is found.

chi square = 3(12-1).9627

chi square = 33(.9627)

chi square = 31.77

Setting alpha equal to 0.05 and entering the appropriate table (Siegel, p 249, table C), the probability of chi square equal to or greater than 31.77 with df = 11 was 0.01 > p > 0.001. Therefore the null hypothesis can be rejected at an alpha value equal to 0.01. Thus a large and highly significant correlation between survey dimensions and time can be concluded.

Next, W was computed for the array of survey indices and waves. In this array k = 3 and N = 24. Summing, Rj yielded

900 therefore $\frac{\Sigma R_j}{N} = \frac{900}{25} = 37.5$. Solving as before, the value of s was found to be s = 10,200, therefore,

$$W = \frac{10200}{1/12(3)^2(24^3-24)}$$

$$W = \frac{10200}{10350}$$

$$W = 0.9855$$

Converting to chi square as before,

chi square = 3(24-1) 0.9855

chi square = 69(0.9855)

chi square = 67.9995 Entering the tables, the probability of a value of chi square equal to or greater than 67.9995 with df = 23 was found to be $p < 0.001$. It could therefore be readily inferred that not only was the null hypothesis rejected but that the correlation between survey indices and time was even stronger than for dimensions.

Taking the value of W computed for the array of indices and substituting it into the linear formula for the average value of Spearman's rho the following results were obtained:

$$r_s(avg) = \frac{KW-1}{K-1} = \frac{(3)(0.9855-1)}{2} \quad 0.9783$$

The same computation for the dimensional variables yielded:

$$r_s(avg) = \frac{(3)(0.9627-1)}{2} = 0.9441$$

It was clear that not only was the correlation high over all three Waves but that the correlation between Waves was also extremely high (both values of rho(averaged) were significant at the 0.01 level).

The values of W were considered highly important in that the extremely high correlation found strongly supported the hypothesis that the movement or trend of the survey scores

were a function of the individual movement of independent component variables. Thus any conclusions drawn about the trend of scores could be stated with confidence as applicable to the entire survey as an integral whole. In addition, the high correlation of the individual (Wave) rankings supported the theory that matched-pair tests between Wave I and Wave III were in error to the extent that the total change in scores were not accurately represented. It was therefore evident that to truly infer the degree of change over the entire time frame under consideration, tests which utilized all of the data simultaneously were necessary.

Two statistical tests were decided upon based on the apparent nature of the data. The Kruskal-Wallis one-way analysis of variance was selected for the case of k independent samples and the Friedman two-way analysis of variance was selected for the case of k related samples.

The Kurskal-Wallis one-way analysis of variance by ranks is an extremely useful test for deciding whether k independent samples are from different populations. As applied to the data on survey scores the question is whether the differences among Waves signify genuine population differences or whether they represent chance variations such as are to be expected among several random samples from the same population. This technique tests the null hypothesis that the three Waves come from the same population or from identical populations with respect to mean scores. In the computation of the Kruskal-Wallis test, each of the N observations are replaced by ranks. That is, all of the scores from all of the Waves are combined and ranked in a single series. The smallest score is ranked 1 with ranks increasing as scores up to the largest score having rank N. As before, the means for survey dimensions and indices are used, therefore for the two tests N = 36 and N = 72

respectively.

If the null hypothesis is true, then H (the test statistic) is distributed as chi square with df = k-1, provided that the sample sizes are not too small. H is defined by the formula,

$$H = \frac{12}{N(N+1)} \Sigma \frac{R_j^2}{N_j} - 3(N+1)$$

where

k = the number of Waves = 3

nj = the number of cases in the j'th Wave

N = the sum of nj's (the number of cases in all samples combined)

Rj = the sum of the ranks in the j'th column

Fig 22 shows the rank distribution of dimension means in the proper format.

Dimension	Wave I	Wave II	Wave III
1	3	1	1
2	28	30	31
3	19	18	21
4	34	33	35
5	29	23	25
6	6	4	5
7	17	12	15
8	10	7	8
9	20	13	14
10	36	24	32
11	11	9	16
12	26	22	27
	$R_1 = 239$	$R_2 = 196$	$R_3 = 231$

Figure 22 - RANK DISTRIBUTION OF DIMENSION MEANS OVER WAVE ARRANGED BY THE KRUSKAL-WALLIS METHOD

Where rank ties occur the average of the ranks for which the scores tied are substituted. Although a correction for ties can be made to the formula for H, the value is rarely changed by more than 10% and the correction tends to inflate the value. Therefore the correction formula was not applied even though some ties existed since it was felt that the effect was negligable.

Utilizing the values of Rj from Figure 22, the computation of H for the dimensional matrix is as follows:

$$H = \frac{12}{36(36+1)}\{(239)^2/12 + (196)^2/12 + (231)^2\} - 3(36+1)$$

$$H = \frac{12}{1332}(12408.166) - 111$$

$$H = 12(9.3154) - 111$$

$$H = 111.785 - 111$$

$$H = 0.785$$

Utilizing the table of probabilities for chi square (Siegel, p 249, table C) with df = 2 and letting alpha equal 0.05, the probability under the null hypothesis that H is greater than or equal to 0.785 was $p > 0.5$. Since $0.5 > 0.05$ failure to reject the null hypothesis is decided. Subject to the assumption that the three waves are independent it was concluded that no significant differences existed between the means of the dimensions over the time interval under consideration.

Fig 23 shows the rank distribution of survey index means and the values of Rj.

Index	Wave I	Wave II	Wave III
1	16	20	22.5
2	7	5	4
3	19	14	18
4	9	6	8
5	3	2	1
6	65	67	72
7	41	43	44
8	68	47	64
9	26	32	36
10	69.5	46	71
11	37	30	35
12	10	11	15
13	40	42	48
14	49	45	52
15	66	69.5	63
16	59	54	61
17	58	50	55
18	17	12	13
19	38	28	33
20	25	21	22.5
21	39	29	31
22	62	53	60
23	27	24	34
24	56	51	57
	$R_1 = 906.5$	$R_2 = 801.5$	$R_3 = 920$

Figure 23 - RANK DISTRIBUTION OF INDEX MEANS OVER WAVE ARRANGED BY THE KRUSKAL-WALLIS METHOD

Proceeding as before with N = 72, nj = 24, and k = 3 H is computed as follows:

$$H = \frac{12}{72(72+1)} \left\{ \frac{(906.5)^2}{24} + \frac{(801.5)^2}{24} + \frac{(920)^2}{24} \right\} - 3(72+1)$$

$$H = 12\{6.5143 + 5.0926 + 6.7098\} - 219$$

$$H = 219.8004 - 219$$

$$H = 0.8004$$

Using the same table as before with df = 2 and alpha equal to 0.05 the probability under the null hypothesis of H = 0.8004 was $p > 0.5$ and since $0.5 > 0.05$ the null hypothesis could not be rejected at the 0.05 level of significance. Again, subject to the assumption that the three Waves are independent samples, the conclusion was made that no significant differences existed between the Waves over time with regard to the mean scores of indices.

In contrast to the Kruskal-Wallis test, the Friedman two-way analysis of variance is used for data when the case is of k related samples. This analysis of variance by ranks is useful for testing the null hypothesis that the k samples (Waves) have been drawn from the same population. It is necessary that the samples be matched in that the number of cases is the same in each sample. For the purposes of this study the matching is achieved by treating the data as three matched sets of scores under three different conditions (time).

For the Friedman test, the scores are cast in a two-way table having N rows and k columns. The rows represent the matched sets of scores and the columns represent the different conditions (Waves). The data of the test are ranks. The scores in each row are ranked separately. That is, with three Waves under consideration, the ranks in any

row range from one to three. The lowest score is given the rank 1 then upwards with the highest score being given rank 3. If the null hypothesis were true then the distribution of ranks in each column would be determined by chance. That is, if the scores were independent of the conditions, the set of ranks in each column would represent a random sample from the discontinuous rectangular distribution and the rank totals for the various columns would be about equal. If the scores were dependent on the conditions (i.e. if the null hypothesis were false), then the rank sums would vary from column to column. The Friedman test determines whether the rank totals (Rj) differ significantly. The test statistic is distributed approximately as chi square with df = k-1 and is defined by the formula

$$\chi^2 = \frac{12}{NK(K+1)} \Sigma R_j^2 - 3N(K+1)$$

where

N = number of rows

k = number of columns

Rj = sum of ranks in the j'th column

No correction for ties can be computed for chi square, Friedman states (Friedman, 1937, p 681) that substitution of the average rank for tied values does not effect the validity of the chi square test.

Fig 24 shows the row rank distribution of dimension means for the Friedman test and the columnar totals (Rj).

Dimension	Wave I	Wave II	Wave III
1	3	1	2
2	1	2	3
3	2	1	3
4	2	1	3
5	3	1	2
6	3	1	2
7	3	1	2
8	3	1	2
9	3	1	2
10	3	1	2
11	2	1	3
12	2	1	3
	$R_1 = 30$	$R_2 = 13$	$R_3 = 29$

Figure 24 - ROW RANK DISTRIBUTION OF DIMENSION MEANS

Proceding from the formula, the computation of χ^2 was as follows:

$$\chi^2 = \frac{12}{(12)(3+1)} \{(30)^2 + (13)^2 + (29)^2\} - 36(4)$$

$$\chi^2 = 1/12(1910) - 144$$

$$\chi^2 = 15.167$$

Using the chi square table (Siegel, p 249, table C) the probability under the null hypothesis that $\chi^2_r = 15.167$ was $p < 0.001$ when df = 3-1 = 2. Since $0.001 < 0.05$ the null hypothesis could be rejected at a value of alpha equal to 0.05.

Fig 25 shows the row rank distribution of index means and the total of column sums (Rj).

Index	Wave I	Wave II	Wave III
1	1	2	3
2	3	2	1
3	3	1	2
4	3	1	2
5	3	2	1
6	1	2	3
7	1	2	3
8	3	2	1
9	1	2	3
10	2	1	3
11	3	1	2
12	1	2	3
13	1	2	3
14	2	1	3
15	2	3	1
16	2	1	3
17	3	1	2
18	3	1	2
19	3	1	2
20	3	1	2
21	3	1	2
22	3	1	2
23	2	1	3
24	2	1	3
	$R_1 = 54$	$R_2 = 35$	$R_3 = 55$

Figure 25 - ROW RANK DISTRIBUTION OF INDEX MEANS

Proceeding with the computation as before,

$$\chi^2_r = \frac{12}{24(3)(3+1)}\{(54)^2+(35)^2+(55)^2\} - 72(4)$$

$$\chi^2_r = 12/288(7166) - 288$$

$$\chi^2_r = 10.58$$

Entering the chi square table with df = 3-1 = 2, the probability under the null hypothesis of $\chi^2_r = 10.58$ was $p < 0.01$. Therefore at alpha equal to 0.05 the null hypothesis was rejected and the conclusion drawn that significant differences existed between Waves with regard to the ranks of index means. It was interesting to note that for index means the null hypothesis could be rejected at the 0.01 level but that for dimension means, which were derived from the index scores, the null hypothesis could be rejected at the 0.001 level of significance. Obviously no significant amount of information was lost when dimension means were compiled from index components.

VI. CONCLUSIONS AND RECOMMENDATIONS

Any conclusion based on the analysis of the HRM data so far presented must be tied to a decision regarding the nature of the samples. Although many of the assumptions inherent in classical statistical analysis were avoided by theuse cf nonparametric techniques, the question of whether the groupings of data, identified as waves, represented related samples or could be treated as independent samples could not be avoided. It should be pointed out that the question of differing results, depending on whether the samples were considered related or independent, in no way violated the reasoning for the selection cf nonparametric methodology. To the contrary, both distribution-free and parametric tests generally require the "sampling assumptions" that sample observations have been drawn randomly and independently from their parent populations, and both are highly vulnerable to violations of this type of assumption. Obviously, the more elaborate the population assumptions the fewer the number of situations which meet (or nearly meet) them, and, in this sense, parametric assumptions are the more susceptible to violation (Bradley, 1968). For example, when the nonparametric assumption of continuous distributions is violated, both the fact and the degree of the violation would tend to be readily apparent from the existence of tied scores in the obtained data. No such indication advises the researcher that a parametric assumption has been violated. On the basis of that arguement, it was consideredthat any violations of basic assumptions made would have resulted in the same degree of inaccuracy, regardless of the methodology.

Since it was strongly felt that no error in methodology existed and that, in fact, all of the statistical tests used were valid given the underlying assumptions a forced choice decision regarding the nature of the samples was required. The decision reached by the authors was that the necessary and sufficient conditions required to satisfy an assumption of independence between the samples (such as random selection) were not met. The conclusion that the samples were related to the extent that independent sample tests were invalid obviously required the rejection of the results obtained from all such tests. The tests in question (the Kolmogorov-Smirnov test and the Kruskal-Wallis test) were included in the thesis as an object lesson for future researchers. Since opposing results occurred (based on tests designed to measure the same factor) dependent on the decision relevant to the relatedness or independence of the samples, any analysis conducted in the future should be responsive to the fact that erroneous conclusions could be drawn from the results of any tests, either parametric or nonparametric, which assume independence of the samples. That is not to say that tests of independent samples cannot be constructed from the HRM data banks. Instead, the point made is that serial observations involving the same organizational units (i.e., non-random sampling) must be treated as related regardless of the fact that the sample membership may have changed greatly over the period of observation.

Based on the decision that the data was representative of related samples the null hypothesis of the thesis was not rejected at the 0.05 level of significance. Friedman's test clearly implied a highly significant change in the ranks of scores over time but visual examination of the data matrix clearly indicated that the change was toward lower ranks in Wave II as compared to Wave I and Wave III. In addition, both the Wilcoxon and Friedman tests clearly indicated that

the dominating change occurred between Wave II and III. Overall the tests involving related samples were very consistent in their results. It can now be said, with a fair degree of confidence, what has occurred relative to HRM survey scores over time and a prediction can be made as to what type of change should occur in the future.

The first and most obvious conclusion to be drawn was addressed earlier in the thesis, which was that it appeared that a curvilinear relationship existed between the survey scores and time. The scores tended to decrease significantly overall between Wave I and Wave II and then tend significantly upward between Wave II and Wave III. It may be recalled that this phenomena was predicted by Likert's work with the SOO, albeit on a more compressed time scale. Assuming that the reason for the initial drop in scores was the same for the HRM survey as for the SOO, that is, a reflection of increased sensitivity to the attitudes addressed in the questionnaire by the respondents, then it could be anticipated that future scores would tend to conform to the general pattern reported by Likert. The prediction would be that scores would tend to continue to increase at a decreasing rate, approaching stability at a median score of approximately 4.0 (on a 5 point scale). The primary question relevant to that prediction involves the amount of time necessary for the scores to stabilize. Due to the high attrition rate in the Navy and the resulting large and constant influx of personnel new to the military organization, it could be anticipated that significant change would take considerably longer than it would in a comparable but more stable civilian environment. The arguement could be made that the data could be sorted on a stability correlated variable such as time in service or paygrade and analysis performed which disregarded the effects of high attrition. Two basic counter-arguements to that proposal are evident; one, the attitudes of lower

paygrade first-term personnel are of extreme importance in that this is precisely the group which the Navy is attempting to attract and retain, two, sorting on paygrade and comparing low with high shows a bimodal distribution with higher paygrades skewed to the left. It is in fact highly possible that the survey scores for E-6s and above (including officers) have already stabilized and that it would be necessary to recalibrate the survey instrument in order to detect any significant future changes in that group.

Assuming the mean score of Wave II to be a base-line figure (i.e. the point at which scores began to trend upward) and further assuming a mean of 4.0 to be a stabilization point, then the amount of change between Wave II and Wave III represented only 4.3% of the anticipated total change. Extrapolating on the basis of that percentage as average for an HRM cycle it would require approximately thirty-three more years for the scores to approach maximum. That is obviously not a realistic prediction but it does indicate the minute degree of change which has occurred to date. The conclusion that positive change has occurred cannot yield the concurrent conclusion that planned change is proceeding swimmingly. To the contrary, the data indicates that while the fact of positive change is encouraging the rate of change and its magnitude have been totally inadequate to indicate any significant change within a realistic framework of time.

There are any number of possible explanations as to why the amount of change has been so small. The attempt to prove or disprove any of them would constitute thesis material for future graduate research in the HRM field. Some of the possible explanations are presented for consideration regardless of the ability to prove or disprove at present.

The most likely explanation involves the nature of the change itself. Using distinctions developed recently (Golembiewski et al, 1976), three basis types of change occur or can occur with psychometric data. The first, alpha change involves a variation in the level of some existential state, given a constantly calibrated measuring instrument related to a constant conceptual domain. It was this type of change which was measured by the tests used in this thesis. Beta change involves a variation in the level of some existential state, complicated by the fact that some intervals of the measurement continuum associated with a conceptual domain have been recalibrated. The concept of beta change is probably best expressed by Upshaw's model (Upshaw, 1962). According to Upshaw, the respondent comes to the questionnaire setting with a psychological range that he imposes upon the object or person specified by the questionnaire. In making a category rating, the respondent divides the psychological range into equal segments of equal size and assigns the object to a category on the basis of the relative distance of the object from the end-points of the psychological range. A beta shift or change specifically refers to a change or shift in the end-points of a particular psychological range. Moreover, a change in the respondent's psychological range should produce a change in the respondent's category assignments. The important aspect of a beta change is that it could have occurred simultaneously with an alpha change. In this case the response scale recalibration would have completely masked the real change in attitude towards the object. The increase in actual change being, for example, offset by the decrease due to response recalibration. Lindell and Drexler (1977) make the argument that the use of multiple items in scales, such as used in the HRM survey, limits the possibility for response scale recalibration since each item is used to direct the respondent's attention to different aspects of the construct that is to be measured. It was not

the purpcse of this thesis to take sides in an intellectual dispute involving the nature of change, only to point out the possibility that the unaccounted for occurance of a beta change may have masked the true extent of actual change.

The third type of change, gamma change, involves a redefinition or reconceptualization of some domain, a major change in the perspective or frame of reference within which phenomena are perceived and classified. Stated very simplisticly as it applies to the HRM survey, a gamma change would involve a change in the perception of the respondents as to what a particular set of questions is referring to or a change in the minds of the respondents as to the definition ofthe object which the questions address. For example, the dimension of Command Climate may have been originally answered by respondents to the questions as referring specifically to the commanding officer. Over time the conception of command may have broadened to include the general command structure of the unit. Obviously, in this case, although the responses are treated statistically as measures of the same variable they may in fact have been responses to completely or significantly different concepts. Golembiewski et al argued convincingly that gamma change can be assessed by investigating the differences in factor structures between two time periods. It is known that a change in factor structure occurred during the period covered by this thesis. In fact, this change in factor structure was instrumental in the revision of the survey and the creation of the form currently in use. Unfortunately the conclusion that a gamma change has occurred cannot be easily made since recent investigation and research has indicated that changes in factor structure can be produced by beta change as well. It is fairly safe to conclude that some change, either beta or gamma or both, has occurred in addition to alpha change. Whether or not that change has masked the actual trend of attitude change and the degree of

effect it has had is obviously material for future research.

Another possible explanation for the small degree of change would involve the possibility that survey respondents have been unconsiously attempting to manipulate the system by indicating attitudes lower on the scale than are actually felt. Lack of familiarity with the form would not have allowed such manipulation initially but subsequent re-surveys could have been amenable to this type of psychological gamesmanship. Similar to the reasoning that if one pill is good for an illness then two would be better, the subconsious realization on the part of survey respondents that low unit scores on a particular dimension resulted in efforts by the command to effect positive change could have resulted in unconscious devaluation of scores in subsequent surveys in an attempt to effect greater efforts by the command. Of course this theoretical phenomena could work in the opposite direction. That is, if it were generally percieved that command efforts were sincere, scores could be over-rated on re-survey (categorizations more positive than actual attitude changewould have been represented) in an unconscious attempt to "reward" the command. Given the extremely large sample size involved and intervening casual variables, it would prove extremely difficult if not impossible to either prove or disprove this theory although it is a consideration which has intuitive merit.

The last explanation considered involved the possibility that the HRM survey was simply not measuring trend in any manner amenable to simple statistical description. The possibility that external variables (i.e. factors not measured by the survey) have had a dominant effect on scores over time cannot be discounted entirely. Research of the literature in the HRM field failed to reveal a case of analysis where correlation of the survey variables and time

was conducted using an external dummy variable. As was previously explained nonparametric tests were chosen for this thesis partially to avoid violation of many of the assumptions of classical techniques. This may have been a case of begging the issue. It could be that an explanation of why certain assumptions would have been violated could have a direct bearing on the statistical results achieved. As Dutta (1975) stated relevant to econometric analysis; given a set of observations taken in successive time periods an assumption is made that a disturbance term (Ut) is included in the relationship between the dependent and independent variables. The further assumption is made that each disturbance term is statistically independent of all other Ut terms associated with different time periods. If the disturbance terms do not fulfill the assumption of mutual independence they are said to be serially correlated. The effect of serial correlation could be to distort the measurement or observation of trend. As an example assume that the unemployment rate in the nation had an indirect effect on the attitudes of individuals towards the Navy. This factor would constitute a disturbance term effecting the scores on HRM surveys. Therefore, since these effects would obviously be serially correlated, if all other factors were held constant the survey scores would tend to fluctuate over time as a function of the unemployment rate. If in this case the distrubance terms were negatively correlated with time the effect would be to depress an otherwise positive trend of the HRM survey variables. Although it might be extremely difficult to identify the nature of serially correlated disturbance terms, tests do exist to determine the existance of serial correlation in a particular case involving the residuals from regression analysis. Since least squares regression was not utilized in this thesis no test for the presence of serial correlation could be conducted but was instead left for future research.

Another problem which could account in part for trend deflection is that of heteroscedasticity (the absence of homoscedasticity). It was beyond the scope of this study to discuss this problem in any detail. The interested reader is directed to Dutta's discussion of the problem (Dutta, 1975). It will suffice to state that in the presence of heteroscedasticity significance tests of the parameters of the analysis are unreliable. It may be recalled that the tests used in this study did not assume homoscedasticity and can therefore be assumed to be more reliable in the presence of heteroscedasticity. This aspect of nonparametric techniques is of additional value in cases such as the HRM survey data where the absolute value of change is extremely small on many of the variables.

Other examples and possible explanations could have been given, but the point should be clear. Human resources are the Navy's most valuable assets. If enlightened decisions are to be made with regard to managing those presious resources, the Navy must be sure that the data on which such decisions are based is the best possible data that can be obtained. The HRM data bank represents a wealth of important information but it must be considered from all possible angles to insure that all pitfalls are avoided. The confidence level of the Navy managers who use this data for decision making must be high therefore the attempt in this thesis to show how different approaches to the same data can result in different conclusions. Much more research in this area is needed if the Navy is to be able to say in the future that the best was done with what was available.

APPENDIX A

THE HRM SURVEY

The Navy Human Resource Management Survey represents an effort to develop a standardized questionnaire which reliably measures the areas of concern defined in the Navy Human Goals Plan. The purpose of the survey is to generate data which will provide each command with quantified information regarding its level of functioning. The datum are to be used to help the command focus on and deal with problems in the areas of organizational development and management, equal opportunity, intercultural relations, training, and drug and alcohol abuse.

The survey contains 88 core questions which are answered on an answer sheet provided to each respondent, and the responses are then processed and summarized by automated equipment. The survey answer sheet also provides spaces for optional or supplementary questions, which individual commands may use to include questions of particular interest for their own work situation. Finally, the survey includes 17 demographic questions (requesting respondent age, race, pay grade, current enlistment status, and the like). These questions are used to see how these characteristics influence answers on the 88 core survey items, or to group data by demographic categories during processing. For example, one of the demographic questions asks for the respondent's supervisor number. A particular number is assigned to all supervisors within a command, and coded on the survey by their subordinates. The number is important

because it allows all data from people in the same work group to be averaged together to form work group level data.

In this section the structure of the HRM Survey, focusing on the way core items are grouped to form indices, and the domain of command or work group functioning which indices and dimensions are designed to measure are described. This structure is outlined below. Each index and its respective questions as stated in the survey are listed. All questions have an answer alternative scale from 1 to 5 as follows:

1 - to a very little extent

2 - to a little extent

3 - to some extent

4 - to a great extent

5 - to a very great extent

Exceptions to this are questions 48 through 53 which are answered using the following scale:

1 - very dissatisfied

2 - somewhat dissatisfied

3 - neither satisfied nor dissatisfied

4 - fairly satisfied

5 - very satisfied

THE HRM SURVEY - DESCRIPTION OF INDICES AND COMPONENT SURVEY ITEMS FROM FORM NINE

COMMAND CLIMATE DIMENSION - refers to the condition, policies, and procedures within which a work group operates. These conditions and policies are created for a work group by other groups, especially by those groups above it in the command hierarchy. Climate conditions set bounds on what does and what does not go on within any work group. Aspects of climate, as listed below, can help or hinder groups, or do both at the same time.

Communication Flow Index - Command leadership understands the work and problems of the command. Information flows freely through the chain of command easily, from the work groups to a listening and responsive leadership and easily to the work groups concerning plans and problems facing the command.

1. Is the amount of information you get about what is going on in other departments or offices adequate tc meet your needs?

2. To what extent are you told what you need to know to do your job in the best possible way?

3. How receptive are those above you to your ideas and suggestions?

Decision Making Index - Information is widely shared within the command and decisions are made at those levels where the most adequate information is available. Supervisors seek out information before making decisions.

4. Decisions are made in this command at those levels

where the most adequate information is available.

5. Information is widely shared in this command so that those who make decisions have access to available know-how.

6. When decisions are being made, to what extent are the people affected asked for their ideas?

Motivation Index - The command motivates personnel to contribute their best efforts through rewards for good performance and career enhancing duties.

7. To what extent do you feel motivated to contribute your best efforts to the command's mission and tasks?

8. Do you regard your duties in this organization as helping your career?

9. Work group members who contribute the most are rewarded the most.

Human Resources Emphasis Index - The command shows concern for human resources in the way it organizes its personnel to achieve its mission. Personnel within the command perceive that the organization and assignment of work sensibly considers the human element.

10. To what extent does this command have a real interest in the welfare and morale of assigned personnel?

11. To what extent are work activities sensibly organized in this command?

12. This command has clear-cut, reasonable goals and objectives that contribute to its mission.

13. I feel that the workload and time factors are adequately considered in planning our work group assignments.

Lower Level Influence Index Lowest level supervisors and non-supervisory personnel have the opportunity to influence what goes on in their departments.

14. In general, how much influence do lowest level supervisors (supervisors of non-supervisory personnel) have on what goes on in their departments?

15. In general, how much influence do non-supervisory personnel have on what goes on in your department?

SUPERVISORY LEADERSHIP DIMENSION - Comprised of the behavior of the supervisor toward his subordinates.

Support Index - Leaders behave in a way which increases the work group member's feelings of worth and dignity.

16. How friendly and easy to approach is your supervisor?

17. When you talk with your supervisor, to what extent does he (she) pay attention to what you are saying?

18. To what extent is your supervisor willing to listen to your problems?

19. My supervisor makes it easy to tell him (her) when things are not going as well as he (she) expects.

Teamwork Index - Supervisors encourage subordinates to develop close, cooperative working relationships with those

who work for them.

20. To what extent does your supervisor encourage the people who work for him (her) to work as a team?

21. To what extent does your supervisor encourage the people who work for him (her) to exchange opinions and ideas?

Goal Emphasis Index - High standards of performance are set, maintained, and encouraged by supervisors.

22. To what extent does your supervisor encourage people to give their best effort?

23. To what extent does your supervisor maintain high personal standards of performance?

Work Facilitation Index - Supervisors help those who work for them to improve performance. Subordinates and supervisors work together to solve problems which hinder task completion and performance.

24. To what extent does your supervisor help you to improve your performance?

25. To what extent does your supervisor provide you with the help you need to you can schedule work ahead of time?

26. To what extent does your supervisor offer new ideas for solving job related problems?

PEER LEADERSHIP DIMENSION - Behavior of work group members toward each other.

Support Index - Work group members behave toward each other in a manner which enhances each member's feeling of personal worth.

27. How friendly and easy to approach are the members of your work group?

28. When you talk with the members in your work group, to what extent do they pay attention to what you are saying?

29. To what extent are the members in your work group willing to listen to your problems?

Teamwork Index - The behavior of work group members encourages the development of close, cooperative wcrking relationships. Work group members maintain and encourage high standards of performance.

30. How much do members of your work group encourage each other tc work as a team?

31. How much do members in your work group stress a team goal?

32. How much do people in your work group enccurage each other to give their best effort?

33. To what extent do people in your work group maintain high standards of performance?

Work Facilitation Index - Work group members help each toher improve performance. The work group workd together to solve problems which hinder performance and task completion.

34. To what extent do members in your work gorup help you find ways to improve your performance?

35. To what extent do members of your work group provide the help you need so you can plan, organize, and schedule work ahead of time?

Problem Solving Index - Work group members work well in solving problems.

36. To what extent do members of your work group offer each other new ideas for solving job related problems?

37. Members of my work group take responsibility for resolving disagreements and working out acceptable solutions.

38. To what extent do people in your work group exchange opinions and ideas?

WORK GROUP PROCESS DIMENSION - Measures those things which characterize the group as a team and whether group members work together well or poorly. The way in which group members share information, make decisions, and solve problems determines the group's productiveness and the quality of its outputs.

Work Group Coordination Index - Work group members plan, coordinate, and support each other effectively.

39. To what extent does your work group plan together and coordinate its effcrts?

40. To what extent do you have confidence and trust in the members cf your work group?

41. To what extent is information about important events widely exchanged within your work group?

42. To what extent does your work group make good decisions and solve problems well?

Work Group Readiness Index - The work group is able to adapt to emergency situations and meet its mission.

43. To what extent has your work group been adequately trained to handle emergency situations?

44. My work group performs well under pressure or in emergency situations.

45. My work group can meet day-to-day mission requirements well.

Work Group Discipline Index - Work group members maintain Navy standards of etiquette and discipline.

46. The members of my work group reflect Navy standards of military courtesy, appearance, and grooming.

47. I feel that Navy standards of order and discipline are maintained within my work group.

ADDITIONAL INDICES FOR EMPHASIS (END RESULTS MEASURES)

Satisfaction Index - Personnel within the command are satisfied with their supervisors, the command, other work group members, their jobs, and their present and future progress in the Navy.

48. All in all, how satisfied are you with the people in your work group?

49. All in all, how satisfied are you with your supervisor?

50. All in all, how satisfied are you with your job?

51. All in all, how satisfied are you with this organization, compared to most others?

52. All in all, how satisfied do you feel with the progress you have made in the Navy, up to now?

53. How satisfied do you feel with the chance for getting ahead in the Navy in the future?

54. Does your assigned work give you pride and feelings of self worth?

Integration of Men and Mission Index - The command is seen as effective in getting people to meet the command's objectives as well as meeting the individual's needs.

55. To what extent is your organization effective in getting you to meet its needs and contribute to its effectiveness?

56. To what extent does your organization do a good job of meeting your needs as an individual?

Training Index - Individuals have been trained in their assigned tasks. The development of technical and leadership skills and other facets of professional advancement are encouraged.

57. I have been adequately trained to perform my assigned task.

58. To what extent has this organization trained you to accept increased leadership responsibilities?

59. To what extent has this organization trained you to accept increased technical responsibility?

General Index - The following questions provide useful data in and of themselves; however, they dc not statistically group with other questions in the Navy Human Resource Management Survey.

60. Our supervisor gives our work group credit for good work.

61. To what extent does your supervisor attempt tc work out conflicts within your work group?

62. People at higher levels of the organization are aware of the problems at my level.

ADDITIONAL INDICES (HUMAN GOALS)

Equal Opportunity (Race Relations) Index - The command ensures equal opportunity for all personnel in such areas as job assignment, advancement, education, rewards and punishment. There is an openess and willingness to address racial issues within the command. NOTE: It should be understood that in addition to these questions other dimensions, such as command climate, indicate the command's ability to effectively manage in order to achieve equal cpportunity.

63. In my chain of command there is a willingness to talk about racial issues.

64. To what extent does this command ensure that you have equal opportunity for advancement in rate (rank)?

65. To what extent does this command ensure that you

have equal opportunity for job assignment?

66. To what extent does this command ensure that you have equal opportunity for housing?

67. To what extent does this command ensure that you have equal opportunity for education and training?

68. To what extent does this command ensure that you receive a fair and objective performance evaluation?

69. To what extent does this command ensure that you have equal opportunity for recreation?

70. To what extent is military justice administered fairly throughout this command?

71. In my chain of command there is a willingness to talk about sex discrimination issues.

72. In this command work assignments are fairly made.

73. People in this command discourage favoritism.

Drug Abuse Index - Personnel in the command have the ability and willingness to recognize and respond appropriately to drug abuse problems.

74. To what extent do you understand the reasons contrubuting to the abuse of drugs?

75. To what extent do members of your work group discourage drug abuse?

76. My supervisor can be depended upon to respond helpfully and appropriately to personnel with drug problems.

Alcoholism Prevention Index - Personnel in the command and supervisors have the ability and willingness to recognize and respond to alcohol problems in an effective and candid manner.

77. To what extent would you feel free to talk to your supervisor about an alcohol problem in your work group?

78. To what extent does this organization promote attitudes of responsibility towards the use of alcoholic beverages?

79. To what extent do members of your work group discourage the abuse of alcoholic beverages?

80. To what extent does this organization provide alternatives to the use of alcohol at social functions?

81. To what extent would your work group accept and support a recovered alcoholic?

Community Interrelationships Index - Personnel are conscious and concerned with the image they project as representatives of the Navy in all locations, and of the United States when overseas. Personnel have been given sufficient training to be able to integrate into the local community and expect to be dealt with fairly in economic transactions. Work group members look forward to visiting foreign countries.

82. Do members of your work group care about the image they project when ashore in this area?

83. Do you consider the effect of your behavior on how people of this area view Navy personnel?

84. To what extent do you expect to be fairly dealt with while spending money in this area?

85. To what extent do you feel you have sufficient understanding of the people and customs of this area to get along in this community?

86. To what extent has information been provided to assist you and (or) your family to adjust to living in this area?

87. Do you have a good understanding of your perscnal role as a representative of the U. S. when overseas?

88. Do members of your work group look forward to visiting foreign countries?

APPENDIX B

PEARSON'S ANALYSIS

Question #	Coefficient	Significance
1	0.0292	0.001
2	0.0439	0.001
3	-0.0088	0.176
4	-0.0268	0.002
5	0/0100	0.146
6	-0/0689	0.001
7	-0.0370	0.001
8	0.0161	0.045
9	0.0113	0.117
10	-0.0778	0.001
11	0.0407	0.001
12	0.0406	0.001
13	-0.0265	0.003
14	-0.0124	0.147
15	-0.0122	0.152
16	0.0457	0.001
17	0.0091	0.167
18	0.0227	0.008
19	0.0375	0.001
20	0.0045	0.317
21	0.0117	0.109
22	-0.0277	0.002
23	0.0101	0.144
24	0.0425	0.001
25	0.0190	0.023
26	0.0498	0.001
27	-0.0087	0.180
28	0.0060	0.264
29	0.0118	0.107
30	0.0025	0.397
31	0.0115	0.113
32	-0.0585	0.001
33	-0.0163	0.044
34	0.0698	0.001
35	0.0156	0.051
36	0.0347	0.001
37	0.0041	0.333
38	0.0530	0.001
39	0.0408	0.001
40	0.0078	0.206
41	0.0350	0.001
42	-0.0015	0.437
43	0.0098	0.154
44	-0.0119	0.108

PEARSON'S ANALYSIS CONTINUED

Question #	Coefficient	Significance
45	0.0254	0.004
46	0.0242	0.006
47	0.0137	0.076
48	0.0116	0.111
49	0.0275	0.002
50	-0.0194	0.020
51	-0.0483	0.001
52	0.0095	0.165
53	0.0258	0.004
54	-0.0341	0.001
55	-0.0055	0.283
56	-0.0249	0.004
57	-0.0430	0.001
58	-0.0256	0.015
59	0.0219	0.032
60	0.0242	0.005
61	-0.0201	0.045
62	-0.0007	0.472
63	-0.0062	0.301
64	-0.0290	0.001
65	-0.0148	0.063
66	-0.0285	0.010
67	-0.0093	0.167
68	-0.0486	0.001
69	-0.0292	0.001
70	-0.0204	0.018
71	-0.0032	0.396
72	-0.0465	0.001
73	-0.0283	0.008
74	0.0093	0.170
75	0.0038	0.376
76	-0.0163	0.085
77	-0.0307	0.449
78	0.0339	0.002
79	0.0177	0.070
80	0.0529	0.001
81	0.0307	0.005
82	-0.0378	0.001
83	-0.0605	0.001
84	0.0529	0.001
85	-0.0395	0.001
86	-0.0363	0.001
87	0.0022	0.415
88	0.1132	0.001

SPEARMAN'S ANALYSIS

Question #	Coefficient	Significance
1	0.0099	0.385
2	0.0433	0.101
3	0.0232	0.247
4	-0.0011	0.487
5	0.0110	0.374
6	-0.0599	0.080
7	-0.0694	0.020
8	-0.0098	0.386
9	-0.0179	0.299
10	-0.0688	0.021
11	0.0699	0.019
12	0.0199	0.279
13	-0.0280	0.204
14	-0.0473	0.133
15	-0.0629	0.070
16	0.0728	0.016
17	0.0590	0.041
18	0.0595	0.039
19	0.0687	0.020
20	0.0283	0.202
21	0.0181	0.297
22	-0.0581	0.043
23	-0.0151	0.328
24	0.0456	0.090
25	0.0109	0.374
26	0.0320	0.173
27	-0.0345	0.155
28	0.0074	0.413
29	0.0188	0.291
30	0.0262	0.221
31	0.0453	0.092
32	-0.0298	0.192
33	-0.0059	0.432
34	0.0424	0.107
35	0.0173	0.306
36	0.0357	0.149
37	-0.0350	0.153
38	0.0740	0.015
39	0.0189	0.290
40	-0.0009	0.490
41	0.0045	0.447
42	-0.0633	0.032
43	-0.0107	0.378
44	-0.0222	0.259

SPEARMAN'S ANALYSIS CONTINUED

Question #	Coefficient	Significance
45	0.0299	0.193
46	0.0155	0.325
47	0.0178	0.301
48	-0.0079	0.409
49	0.0421	0.107
50	-0.0587	0.042
51	-0.0400	0.124
52	0.0184	0.298
53	0.0440	0.104
54	-0.1038	0.001
55	0.0250	0.231
56	-0.0728	0.016
57	-0.0634	0.031
58	0.0085	0.420
59	0.0413	0.165
60	0.0575	0.045
61	0.0608	0.077
62	-0.0190	0.288
63	-0.0033	0.469
64	-0.0110	0.376
65	0.0090	0.397
66	0.0557	0.101
67	-0.0053	0.440
68	0.0290	0.249
69	-0.0401	0.125
70	-0.0135	0.349
71	-0.0038	0.465
72	-0.0365	0.195
73	-0.0095	0.412
74	-0.0200	0.283
75	0.0259	0.273
76	-0.0305	0.238
77	-0.0103	0.383
78	0.0445	0.149
79	0.0000	0.500
80	0.0895	0.018
81	0.0853	0.023
82	-0.0125	0.366
83	-0.0292	0.213
84	0.0372	0.155
85	-0.0201	0.292
86	-0.0556	0.067
87	0.0235	0.260
88	0.1715	0.001

LIST OF REFERENCES

1. Babbie, Earl R., Survey Research Methods, p. 49, Wadsworth Publishing Company, Inc., 1973.

2. Institute for Social Research, Military Manpower and Modern Values, Bowers, D. G. and Backman, J. G. , October 1974.

3. Institute for Social Research, Technical Report, Survey-Guided Development: Data Based Organizational Change, Bowers, D. G. and Franklin, J. I., June 1975.

4. Bradley, James V., Distribution-free Statistical Tests, Prentice-Hall, 1968.

5. Navy Personnel Research and Development Center, TR 76-5, Human resources management and nonjudicial punishment rates on Navy ships, Crawford, K. S. and Thomas, E. D. , August 1975.

6. Institute for Social Research, Navy retention rates and human resources management, Drexler, J. A. and Bowers, D. G., May 1973.

7. Institute for Social Research, The Human Resourse Management Survey: an item analysis, Drexler, J.A., July 1974.

8. Dutta, M., Econometric Metnods, South-Western Publishing Co., 1975.

9. Franklin, J. L. A path analytic approach to describing causal relationships among social-psycholcgical variables in multi-level organizations, Institute for

Social Research, 1973.

10. Institute for Social Research, Hierarchial differences in Navy functioning, Franklin, J. L., February 1974.

11. Golembiewski, R. T., Billingsley, K. and Yeager, Measuring Change and Persistence in Human Affairs; Types of change generated by OD designs, Journal of Applied Behaviorial Science, v. 28, p. 747-711, 1968.

12. Kirk, Roger E., Statistical Issues: A Reader for the Behavioral Sciences, Wadsworth Publishing Co., 1972.

13. Likert, R. L., New patterns of management, McGraw-Hill, 1961.

14. Likert, R. L., The human organization: Its management and values, McGraw-Hill, 1967.

15. Lindell, M. K. and Drexler, J. A., Issues in Using Survey Methods for Measuring Organizational Change. Battelle Human Affairs Research Centers, 1977.

16. Mager, Robert F., Developing Attitude Toward Learning, Fearon-Pitman Publishers, Inc., 1968.

17. Siegel, Sidney, Nonparametric Statistics For The Behavioral Sciences, McGraw-Hill, 1956.

18. Taguiri, R., and Litwin, G. H. (Eds.) Organizational climate, Harvard University Press, 1968.

19. Taylor, J. C. and Bowers, D. G., Survey of organizations: A machine-scored standardized questionnaire instrument, Institute for Social Research, 1972.

20. Upshaw, H. S., Own attitude as an anchor in equal appearing intervals, Journal of Abnormal and Social Psychology, v. 69, p. 85-96, 1962.

21. Varney, G. H., Organizational Development for Managers,

Addison-Wesley, 1977.

22. Warwick, D. P. and Lininger, C. A., The Sample Survey: Theory and Practice, McGraw-Hill, 1975.

23. Navy Personnel Research and Development Center, Technical Report TN 77-2, The Human Resource Management Survey: A Factor Analysis of each component dimension, Wilcove, Gerry, December 1976.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Professor Carl R. Jones, Code 54Js Chairman, Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
4.	CDR C. B. Gustafson USN, Code 54Gs Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
5.	LCDR Raymond L. Forbes USN, Code 54Fb Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	CDR R. A. McGonigal USN, Code 54Mb Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
7.	LT William L. Carter USN, 2 Gunn Circle Pensacola, Florida 32506	1

8. LT Lawrence L. Yeatman USN, 1
P O Box 129
Naples, Texas 75568

9. Chief of Naval Personnel (Pers-6) 1
Navy Department
Washington, D. C. 20370

10. Commanding Officer Attn. Code 307 1
Navy Personnel Research and Development Center
San Diego, California 92152

11. CAPT James Spurgeon USAF, 1
Route 1, Box 179A
Coolidge, Texas 76635

Thesis 176518
C27385 Carter
c.1 A non-parametric analysis of the HRM data.